

Frederick Douglas
(Admitted Pro Hac Vice)
FEDERAL EXPRESS CORPORATION
3620 Hacks Cross Road
Building B, 3rd Floor
Memphis, Tennessee, 38125
Telephone: (901) 434-8519
Facsimile: (901) 434-9271

SEYFARTH SHAW LLP
Gilmore F. Diekmann, Jr. (SBN: 050400)
Francis J. Ortman, III (SBN: 213202)
560 Mission Street, Suite 3100
San Francisco, California 94105
Telephone: (415) 397-2823
Facsimile: (415) 397-8549

Evelyn L. Becker (CA SBN: 170903)
O'MELVENY & MYERS LLP
1625 Eye Street, NW
Washington, DC 20006
Telephone: (202) 383-5300
Facsimile: (202) 383-5414

Attorneys for Defendant
FEDEX EXPRESS

UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF CALIFORNIA

DERRICK SATCHELL, KALINI BOYKIN,
VALERIE BROWN, RICK GONZALES,
CYNTHIA GUERRERO, RACHEL
HUTCHINS, DANIEL SHERMAN, KELVIN
SMITH, SR., and KEN STEVENSON, on
behalf of themselves and all others similarly
situated,

Plaintiff,

v.

FEDEX EXPRESS, a Delaware corporation,

Defendant.

Case Nos. C03-02659 SI; C03-02878 SI

**DEFENDANT'S OPPOSITION TO
PLAINTIFFS' MOTION IN LIMINE
NO. 4 TO PRECLUDE DEFENDANT'S
EXPERT MARY BAKER'S ANALYSIS
AND OPINION CONTROLLING FOR
THE RACE OF MANAGERS**

DATE: March 2, 2007

TIME: 9:00 a.m.

CTRM: The Hon. Susan Illston

TABLE OF CONTENTS

		<u>Page</u>
1	I. PRELIMINARY STATEMENT	1
2	II. ARGUMENT	3
3	A. The Jury Is Entitled To Hear Dr. Baker's Statistical Analysis And	
4	Conclusions In Determining Whether Systemic Bias Is Operating At	
	FedEx	3
5	B. Dr. Baker's Statistical Analysis And Conclusions Are Wholly Reliable	9
6	III. CONCLUSION.....	11
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26	MOTION IN LIMINE NO. 4 TO PRECLUDE DEFENDANT'S EXPERT MARY BAKER'S ANALYSIS AND OPINION CONTROLLING	
27	FOR THE RACE OF MANAGERS	
	Case Nos. 03-2659 and 03-2878	

TABLE OF AUTHORITIES**Page****Cases**

<i>Anderson v. Anheuser-Busch, Inc.</i> , 65 F. Supp. 2d 218 (S.D.N.Y. 1999)	5
<i>Castaneda v. Partida</i> , 430 U.S. 482 (1977).....	4
<i>Elrod v. Sears, Roebuck & Co.</i> , 939 F.2d 1466 (11th Cir. 1991)	4
<i>Fresenius Med. Care Holdings, Inc. v. Baxter Int'l, Inc.</i> , No. C 03-01431 SBA(EDL), 2006 WL 1646113 (N.D. Cal. June 12, 2006)	9
<i>Goberman v. Washington County</i> , 2001 U.S. Dist. LEXIS 7770 n. 6 (D. Or. April 26, 2001)	4
<i>Holston v. Sports Authority, Inc.</i> , 136 F.Supp.2d 1319 (N.D. Ga. 2000)	4
<i>Leicht v. Hawaiian Airlines, Inc.</i> , 77 F. Supp. 2d 1134 (D. Haw. 1999).....	4
<i>Marlow v. Office of Court Admin.</i> , 820 F. Supp. 753 (S.D.N.Y. 1993)	5
<i>Oncale v. Sundowner Services</i> , 523 U.S. 75 (1998).....	4
<i>Pesok v. Hebrew Union College-Jewish Inst. of Religion</i> , 235 F. Supp. 2d 281 n.5 (S.D.N.Y. 2002)	4
<i>Toliver v. Cmty. Action Comm'n To Help the Economy, Inc.</i> , 613 F. Supp. 1070 (S.D.N.Y. 1985)	4, 5
<i>Welch v. Delta Air Lines, Inc.</i> , 978 F. Supp. 1133 (N.D. Ga. 1997)	5

Other Authorities

K. Kraiger and J.K. Ford, "A Meta-Analysis of Race Effects in Performance Ratings," <i>Journal of Applied Psychology</i> , Vol. 70, 1985, p. 56-65)	6, 11
--	-------

I. PRELIMINARY STATEMENT

Plaintiffs' attempt to preclude the jury from knowing the race of the manager evaluating hourly class members is the ultimate in hypocrisy. *First*, those African American managers are themselves class members intending to bring before the jury claims that they were undervalued compared to their contributions as managers. At the same time, they now appear to allege that they themselves made discriminatory decisions about those they supervise. *Second*, the named plaintiffs have at every turn helpful to their case, pointed the finger at white managers. Despite the anecdotal evidence they have attempted to gather that white managers discriminate, they no appear to favor of a theory that African American managers throughout the Western Region are engaging in systematic discrimination against African American hourly employees. *Third*, their own experts have relied heavily on studies theorizing that white managers are more likely than minority managers to discriminate on the basis of race. By their present motion, they now appear to contradict the very studies relied upon by their own experts to support their claims of discrimination. The only aspect of plaintiffs' motion that is not ironic is their desire to keep from the jury incredibly powerful evidence that the reason that minorities, on average, receive slightly lower performance evaluation scores and more discipline is because of actual differences between the performance of minority and non-minority employees rather than any racial discrimination by managers.

Beyond the hypocrisy, however, plaintiffs' entire motion rests upon the legal argument that the jury is not entitled to take in to account the race of the manager when determining whether the evidence supports an inference of discrimination. That is simply not the law. There is a wealth of case law holding unequivocally that the jury *is* entitled to consider evidence of the whether the decisionmaker is in the same protected class as the affected employee in determining whether discrimination occurred. Moreover, there are numerous scientific studies – including studies conducted by and relied upon by plaintiffs' own experts – that demonstrate that minority decisionmakers are less likely to demonstrate bias towards minorities than Caucasian decisionmakers. Dr. Baker's analysis demonstrates that, at Fed Ex, managers of all races are

1 similarly likely to issue lower performance ratings and discipline to minority employees,
2 supporting the inference that discrimination is not the basis for the disparity. Indeed, if it were
3 true that (1) FedEx's discipline and performance evaluation processes are fair, and (2) minorities
4 employees at FedEx are more likely to commit infractions than Caucasian employees, one would
5 expect exactly the results that Dr. Baker sees – that managers of all races issue greater discipline
6 and lower performance ratings to minority employees. (See June 30, 2006 Baker Report at 9-14,
7 16-17 ("Baker June Report"); December 20, 2006 Baker Report ("Baker December Report") at
8 16-23, 25-28.) And, Dr. Baker's statistical evidence also shows that there *are* true differences in
9 minority and Caucasian absence and tardiness rates, with minorities significantly more absent or
10 tardy than expected in proportion to their representation in the FedEx workforce. (Baker
11 December Report at 17-28.) These data, in addition to the manager race data, all support the
12 inference that any racial disparities in discipline or performance evaluation rates at FedEx are
13 due to true differences in minority performance, and not discrimination. Moreover, these data
14 also support the inference that criteria that plaintiffs' expert classifies as "tainted" in fact are not
15 tainted at all. The jury is entitled to hear the totality of this evidence.

16 Not only have plaintiffs failed to show that Dr. Baker's analysis is irrelevant, they also
17 have failed to credibly question the reliability of Dr. Baker's analysis. Plaintiffs raise no
18 question as to the statistical methodology employed by Dr. Baker, her qualifications to perform
19 such analyses, nor the mathematical computations or results she presents. Moreover, analyses
20 controlling for the race of the decisionmaker in assessing whether or not discrimination has
21 occurred are so commonsense and common in professionally-refereed literature, including
22 literature relied upon by plaintiffs' own expert, that plaintiffs' claim that the methodology was
23 "created" for this litigation amounts to nothing more than nonsense.

II. ARGUMENT

A. THE JURY IS ENTITLED TO HEAR DR. BAKER'S STATISTICAL ANALYSIS AND CONCLUSIONS IN DETERMINING WHETHER SYSTEMIC BIAS IS OPERATING AT FEDEX.

Plaintiffs seek to shield the jury from knowing the race of the managers who are issuing performance evaluations at FedEx. Specifically, they don't want the jury to know that members of the African American Management Sub-Class are similar to their Caucasian counterparts in that they rate minorities on average slightly lower in performance evaluations. This evidence is relevant to a determination of whether it is discrimination or something else, such as true differences in performance, that has led minorities to receive differing average numbers of discipline and average performance evaluation scores at FedEx.

Dr. Baker begins her analysis with underlying statistical evidence that minorities receive more discipline than expected given their proportion of the FedEx workforce, and receive slightly lower group mean performance scores than Caucasian employees. (Baker June Report at 9, 11, 13, 16; Baker December Report at 16, 20, 24, 27). Dr. Baker took this underlying data and controlled for the race of the manager making the decisions regarding discipline or performance. The analysis Dr. Baker reported confirmed the same pattern existed for *all* racial groups of managers. (*See, e.g.*, Baker June Report at 10-11, 13-14, 16-17, Tables AD-2, AD-3, AD-5, LD-2, LD-3, AP-2, AP-3, LP-2, LP-3; Baker December Report at 17-18, 20-21, 25-27, Tables 6A, 7A, 9A, 13A, 15A, 20A, 21A, 22A, 23A, 29A, 6L, 7L, 9L, 20L, 21L, 22L, 23L; January 10, 2007 Baker Report at 4-5, 7, 9.)¹

In a race discrimination case, the jury is entitled to know the race of the decisionmaker. *No* case known to FedEx, and *no* case cited by plaintiffs, has held that the racial identity of an alleged discriminator, or of a person allegedly involved in establishing an environment hostile to

¹ Plaintiffs cite to a number of pages in Dr. Baker's reports as containing the analysis and opinions they now seek to preclude. (*See, e.g.*, Pls' Motion at 3-5.) These page ranges also contain additional analysis and opinions that are not subject to plaintiffs' motion.

minorities, is irrelevant to evaluating a claim of discrimination.² To the contrary, legions of cases have rejected that principle, holding that the protected status of the manager *is relevant* to deciding whether an inference of discrimination may be established. *See, e.g., Elrod v. Sears, Roebuck & Co.*, 939 F.2d 1466, 1471 (11th Cir. 1991) (no “substantially probative evidence” of discrimination where decision was made by members of the same protected group as plaintiffs); *Toliver v. Cmty. Action Comm’n To Help the Economy, Inc.*, 613 F. Supp. 1070, 1074 (S.D.N.Y. 1985), *aff’d*, 800 F.2d 1128 (2d Cir. 1986), *cert denied*, 479 U.S. 863 (1986) (“given the racial and sexual composition of the [employers’] Board, the inference of discrimination which plaintiff seeks to have this Court draw... is, to say the least, attenuated.”); *Pesok v. Hebrew Union College-Jewish Inst. of Religion*, 235 F. Supp. 2d 281, 287 n.5 (S.D.N.Y. 2002) (“Notably, both Lutwak and Panken are Caucasian and Jewish themselves, making an inference of discrimination more difficult.”); *Goberman v. Washington County*, 2001 U.S. Dist. LEXIS 7770, at *25 n. 6 (D. Or. April 26, 2001) (“When the decisionmaker is close in age to the plaintiff and is also within the protected category, the inference of discrimination is diminished” (citations omitted)); *Holston v. Sports Authority, Inc.*, 136 F.Supp.2d 1319, 1335 (N.D. Ga. 2000) (noting that the fact that the decision-maker and the employee both were Black bolsters a conclusion of non-discrimination, and explaining, “[W]hen the decision makers are in the same protected class as the employee complaining about an adverse employment decision, the employee faces a more difficult burden in establishing that a discriminatory animus played a role in the decision complained about.”); *Leicht v. Hawaiian Airlines, Inc.*, 77 F. Supp. 2d 1134, 1145 (D. Haw. 1999), *rev’d on other grounds* 2001 U.S. App. LEXIS 18016 (9th Cir. 2001) (“When

² *Oncale v. Sundowner Services*, 523 U.S. 75 (1998), cited by plaintiffs, merely rejected “any *conclusive presumption* that an employer will not discriminate against members of his own race.” *Id.* at 78 (emphasis added). It did *not* reject the concept that such an act is unlikely, and especially did not rule that the race of the employer is *irrelevant*. There is an enormous logical gap between the absence of a *conclusive presumption* of same race non-discrimination and an assertion that the racial identity of an allegedly discriminatory employer is irrelevant, as plaintiffs contend here. A finding of the former does not even *hint* at the latter. *Castaneda v. Partida*, 430 U.S. 482, 499 (1977) likewise only rejected a “presum[ption] as a matter of law” of same-race non-discrimination: it did *not* rule, or even suggest, that the race of the person alleged to have discriminated against a group is irrelevant or immaterial.

1 the decisionmaker is close in age to the plaintiff and is also within the protected category, the
 2 inference of discrimination is diminished"); *Anderson v. Anheuser-Busch, Inc.*, 65 F. Supp. 2d
 3 218, 229 (S.D.N.Y. 1999) ("That Williams was a member of Anderson's protected
 4 class...further weakens the inference of discrimination"); *Welch v. Delta Air Lines, Inc.*, 978 F.
 5 Supp. 1133, 1153 (N.D. Ga. 1997) ([I]t is extremely difficult for a plaintiff to establish
 6 discrimination where the allegedly discriminatory decision-makers are within the same protected
 7 class as the plaintiff."); *Marlow v. Office of Court Admin.*, 820 F. Supp. 753, 757 (S.D.N.Y.
 8 1993), *aff'd*, 22 F.3d 1091 (2d Cir.) (indicating that "if [the] decision-maker is in the same
 9 protected class as plaintiff, claims of discrimination become less plausible," *citing Toliver*, 613
 10 F. Supp. at 1074).

11 In fact, plaintiffs' own experts have presented evidence that supports the relevance of Dr.
 12 Baker's analysis controlling for the race of the decisionmaker. In a published law review article,
 13 Dr. Anthony Greenwald presented evidence that African Americans display explicit bias that
 14 either favors or is neutral toward other African-Americans over 95% of the time, whereas
 15 European-Americans display explicit bias that either favors or is neutral to African-Americans
 16 59% of the time. African-Americans also display implicit bias that either favor or is neutral
 17 toward other African-Americans approximately 68% of the time, whereas the same number for
 18 European-Americans tested was 29%. (September 3, 2006 Greenwald Report ("Greenwald
 19 Report") (Ex. A to Declaration of Evelyn Becker, filed herewith) Attachment 3, at 10-12, Table
 20 2; October 27, 2006 Greenwald Deposition ("Greenwald Depo") (Ex. B to Declaration of Evelyn
 21 Becker, filed herewith) at 460-61.) Accordingly, Dr. Greenwald's study shows that African-
 22 Americans are *far less* prone than European-Americans to exercise bias against African-
 23 Americans. (*See, e.g.*, Greenwald Depo. at 89-90 ("If you wanted to select a jury that you could
 24 be sure was lacking in net implicit bias one way or the other towards white and black, our data
 25 shows the best you could do is select a 100 percent black jury"); November 13, 2006 Bielby
 26 Deposition (Ex. C to Declaration of Evelyn Becker, filed herewith) at 447 (agreeing that
 27 Greenwald's study "suggests that African Americans are, in general, less likely to be biased
 28

1 against African Americans than non-African Americans,” and are “more likely to be biased in
2 favor of African Americans”).

3 Plaintiffs’ expert Dr. William Bielby also relies in his report on a 1985 article by Kraiger
4 and Ford that demonstrates that same-race decisionmakers tend to favor ratees of their own race.
5 The Kraiger and Ford study was a meta-analysis of 84 studies of performance review systems. It
6 was this meta-analysis that included evidence “suggesting that white raters give significantly
7 higher ratings to whites than to African Americans; and (2) that African American raters give
8 significantly higher ratings to African Americans than do whites.” (Bielby Report (Ex. D to
9 Declaration of Evelyn Becker, filed herewith) at 10 (*citing* K. Kraiger and J.K. Ford, “A Meta-
10 Analysis of Race Effects in Performance Ratings,” *Journal of Applied Psychology*, Vol. 70,
11 1985, p. 56-65).)

12 While FedEx does not necessarily agree with the conclusions drawn by Drs. Greenwald
13 and Bielby, plaintiffs, having relied on such studies, may not now be heard to complain about
14 evidence that shows that at FedEx, minorities and Caucasians are equally likely to determine that
15 minorities performed less well, supporting an inference that something other than bias is
16 operating at FedEx.

17 The case law and plaintiffs’ own experts’ reports and testimony demonstrate that
18 consideration of the race of the decision-maker, while not necessarily dispositive, is relevant to
19 the jury’s consideration of a finding of discrimination *vel non*. Dr. Baker’s statistical analysis
20 showing that managers of all races similarly issue more discipline and lower ratings to minorities
21 provides the jury with just that relevant information. What conclusions, if any, the jury chooses
22 to draw from this statistical information is of the jury’s choosing, but it should not be deprived of
23 its right to consider this relevant evidence.

24 Not only do plaintiffs challenge the admission of Dr. Baker’s statistical evidence
25 controlling for the race of the manager, they also challenge the conclusions she draws from that
26 evidence. As Dr. Baker has testified, it is well-accepted that statistical disparities, such as the
27 statistical disparities in the discipline levels and performance scores between minority and
28

1 Caucasian employees at FedEx, show no more than that the differences are not likely to have
2 occurred by chance. (*See, e.g.*, Baker December Report at 7.) Discrimination is one potential
3 explanation for the statistical disparities. (*See, e.g.*, Baker December Report (Ex. E to
4 Declaration of Evelyn Becker, filed herewith) at 7-8.) There is, however, another explanation for
5 the disparities that plaintiffs are attempting to keep from the jury – that the disparities could be
6 the result of actual “racial or ethnic differences in a factor that influences the variable of interest
7 that happens to be correlated with race or ethnicity and for which the analysis did not account.”
8 (*See, e.g.*, Baker December Report at 8). To test whether actual differences correlated with race
9 or ethnicity impacted the discipline and performance evaluation statistical results, one should
10 consider the impact of the omitted variable – in this case, race of the managers – on the outcome.
11 (*Id.*)

12 Dr. Baker statistical analysis controlled for one such omitted variable – race of the
13 manager making the decision regarding discipline or performance. Contrary to plaintiffs’ bald
14 assertion, this control test was *not* based on the assumption that minority managers do not
15 discriminate against other minorities. Indeed, as plaintiffs agree, Dr. Baker never stated such an
16 assumption in her reports (Pls’ Motion at 1-2), nor does she intend to offer such an opinion at
17 trial. That is a matter for other plaintiff and defense experts in this case to address. As Dr. Baker
18 did state in her reports, her conclusion was based on the assumption that if the discipline and
19 performance review processes are fair, and true differences between minority and Caucasian
20 performance or infraction rates do exist, then it would follow that managers of all races would
21 issue similar levels of discipline and similar performance evaluation scores. (*See, e.g.*, Baker
22 December Report at 16-17.) Of note, and contrary to plaintiffs’ claim, at no point does Dr.
23 Baker state or assume that “statistical disparities *must* reflect actual performance and disciplinary
24 differences if similar rates are found for managers of different races.” (Pls’ Motion at 8)
25 (emphasis added). Rather, Dr. Baker, consistently with the case law and testimony of plaintiffs’
26 experts, presented her analysis as a means of demonstrating that actual performance and
27 disciplinary differences *may* be the underlying reason for the statistical disparities if such
28

1 similarities are found. (*See, e.g.*, Baker December Report at 16 (“Plaintiffs fail to consider that
2 the racial difference between discipline rates *may* be attributable to true racial differences in the
3 propensity to commit infractions for which employees are disciplined” (emphasis added).) In
4 sum, Dr. Baker’s analysis can be interpreted to demonstrate that true differences correlated with
5 race or ethnicity are one plausible explanation of the disparities – an interpretation of the
6 evidence that the jury is entitled to consider.

7 Dr. Baker’s conclusion that true differences in performance appear to be attributable to
8 different levels of performance and attendance among racial groups is all the more relevant
9 because it is also supported by actual record evidence of the rates of infractions by those groups.
10 As Dr. Baker reported, both African-American and Latino employees at FedEx are tardy or
11 absent from work significantly more often than their Caucasian counterparts. (Baker December
12 Report at 18-19, 21-22.) True to Dr. Baker’s conclusion, “managers of all races disciplined
13 African-Americans for attendance and punctuality infractions at a higher rate than Caucasian
14 employees,” (Baker December Report at 19, 22-23), just as one would expect based on the
15 objective data showing substantially different infraction rates. (Baker December Report at 25-
16 27.)

17 Plaintiffs’ attempts to keep the jury from knowing that the disparity in Latino discipline is
18 the result of decisions made by African American managers who are members of the
19 management class is similarly unavailing. This evidence is part of a larger picture of statistical
20 evidence that shows that the disparity is not the result of a pattern and practice of discrimination.
21 Dr. Baker’s analysis demonstrated that the only significant disparity between actual and expected
22 disciplinary actions against Latino employees were at AGFS, and that the disparity is the result
23 of decisions made by African American AGFS managers who comprise the management class,
24 as opposed to Latino or Caucasian managers. (Baker June Report at 11-12, Tables LD-1, LD-2;
25 Baker December Report at 20-21, Tables 4L, 6L.) Dr. Baker went on to show that the excess
26 Latino disciplinary actions at AGFS issued by African American managers were entirely
27 attributable to the greater number of attendance and punctuality violations engaged in by Latinos

as opposed to Caucasians. (Baker December Report at Table 8L.) Therefore, the data demonstrated that the greater number of disciplinary actions against Latinos at AGFS were due to African American managers disciplining for actual attendance and punctuality violations. This data is relevant to demonstrate to the jury that the disparity in disciplinary actions against Latinos at AGFS is due, not to a systematic culture of discrimination to which all FedEx managers subscribe, but to actual performance differences enforced by African American managers, who are themselves class members. Even if the African American managers presenting their own claims of discrimination intend to testify that they themselves engaged in discrimination against Latinos, FedEx is still entitled to rebut the plaintiffs' argument that this is part of a broader pattern of discrimination by all managers.

Given plaintiffs' allegation that differences in region-wide average discipline rates and performance review scores demonstrate "bias" against minorities, it is entirely relevant for FedEx to present Dr. Baker's conclusion drawn from her statistical analysis controlling for manager race and actual objective performance data that it is true differences in performance – and not some form of infectious bias – that resulted in the disparities in minority and Caucasian discipline and group mean performance scores. To preclude FedEx from presenting such relevant evidence to defend against one of the central theories of plaintiffs' case, and from responding to the evidence plaintiffs intend to present in support of that theory, would unfairly and inappropriately prejudice FedEx.³

B. DR. BAKER'S STATISTICAL ANALYSIS AND CONCLUSIONS ARE WHOLLY RELIABLE.

In their motion, plaintiffs do not dispute the reliability of the actual statistical analysis of the data that Dr. Baker undertook. The facts and data on which Dr. Baker relied are the same facts and data that plaintiff's statistical expert, Dr. Richard Drogin, relied upon in conducting his

³ Because the evidence is so relevant to the issues in this case, plaintiffs' argument that its introduction would be more prejudicial than probative must also fail. As this Court has observed, in applying Fed. R. Evid. 403's balancing test, "the general rule is that the balance should be struck in favor of admission" and where "evidence is relevant to a central issue in a case, admission is warranted." *Fresenius Med. Care Holdings, Inc. v. Baxter Int'l, Inc.*, No. C 03-01431 SBA(EDL), 2006 WL 1646113, at *3 (N.D. Cal. June 12, 2006) (citation omitted).

analyses. At no time do plaintiffs assert that the commonly-used statistical method by which Dr. Baker has controlled for the variable of the race of managers is incorrect or unreliable. Indeed, plaintiffs could not do so because their own statistical expert also controlled for various factors in the same fashion when performing his analyses of FedEx data. In fact, none of plaintiffs' experts ever has questioned the mathematical computations nor the numbers Dr. Baker produced as a result of her analysis.⁴ As a labor economist and statistician with twenty years of experience, who has been qualified by numerous courts in conducting statistical analyses of the outcome of employment decisionmaking processes, Dr. Baker's qualifications to perform this analysis controlling for the race of the managers are above reproach. Because there is no question that the actual statistical data Dr. Baker has produced is reliable, the data and statistical analysis cannot be excluded on the basis of reliability.

Plaintiffs do argue, however, that Dr. Baker's opinion that controlling for the race of the manager may demonstrate whether true performance and infraction differences (as opposed to discrimination) underlie the statistical disparities, is untested and "created for this litigation." (Pls' Motion at 7-8.) That assertion, however, is nonsense. Dr. Baker's analysis controlling for the race of the manager is, in reality, based on a tried and true statistical method. It is without question that a significant disparity among the group mean discipline rates and group mean performance scores of minority and non-minority employees demonstrates nothing more than that the disparity did not occur by chance. To test those disparities to determine the plausible explanations for them is similarly consistent with basic statistical analysis of employment decisions. Given the clear relevance of race in a discrimination case, it is all the more absurd to complain that an analysis that controls for the race of the manager to determine whether

⁴ Plaintiffs cite to this Court's order in which it did not require plaintiffs to admit or deny that there is no statistically significant difference in the frequency with which managers of various races impose discipline on minority employees as record evidence of the unreliability of Dr. Baker's methods. (Pls' Motion at 9.) While the Court ruled that plaintiffs were not obligated to respond further to the request for admissions, its ruling was not based on the lack of inherent reliability of the statistical evidence, but on plaintiffs' questioning of the weight of the data due to anecdotal evidence plaintiffs contend show that certain African American managers were co-opted into discrimination. In other words, while plaintiffs can present the jury with evidence challenging the weight of the statistical evidence, such anecdotal evidence does not render the statistical evidence itself unreliable or inadmissible.

1 discrimination or something else has occurred is unreliable. Even plaintiffs' own expert, Dr.
 2 Bielby, relies on a meta-analysis of a number of studies in professionally-refereed literature in
 3 which performance ratings were analyzed by controlling for the race of the rater to determine the
 4 reliability of the evaluation process being tested. (Bielby Report at 10 (*citing* Kraiger and Ford
 5 at 56-65).) Indeed, the Kraiger and Ford analysis of studies controlling for the race of the rater
 6 concluded that there *are* differences in how Caucasian and African-American raters rate ratees of
 7 their own race – indicating some level of bias in the ratings. (Kraiger & Ford (Ex. F to
 8 Declaration of Evelyn Becker, filed herewith) at 61-62.) Thus, plaintiffs' argument that Dr.
 9 Baker "created" a methodology – a methodology based in basic statistical analysis controlling
 10 for the race of the rater to determine if some factor other than bias caused the observed statistical
 11 disparities – is simply groundless.

12 Plaintiffs also argue that the theory that statistical disparities may reflect true differences
 13 in performance if, controlling for the race of the managers, all managers regardless of race issue
 14 greater discipline and lower group mean performance scores to minorities is based on an
 15 unfounded and "unspoken principle that African-American and Latino managers will not
 16 discriminate against African-American and Latino employees." (Pls' Motion at 7.) This
 17 assertion is yet another example of plaintiffs' propensity to mischaracterize an argument in order
 18 to refute it when they cannot refute the actual point being made. As explained in Section I
 19 above, at no time does Dr. Baker assert that minorities never discriminate against minorities, nor
 20 is that an opinion she intends to give at trial. Therefore, plaintiffs' questioning of Dr. Baker's
 21 qualifications to conduct an analysis on that basis is a red herring. Nonetheless, the assumption
 22 that African-Americans are less likely to discriminate against African-Americans is supported by
 23 testimony of *plaintiffs' own experts*, studies relied upon by those experts, and is a factor that
 24 numerous courts have held is relevant to a determination of discrimination.

25 III. CONCLUSION

26 For all the foregoing reasons, plaintiffs' motion to preclude Dr. Baker's analysis and
 27 opinion controlling for the race of managers should be denied in its entirety.

FedEx reserves the right, absent a Court order to the contrary, to file a response to plaintiffs' reply brief, if any, should it contain new, revised or different arguments or factual assertions not contained in plaintiffs' motion in limine.

Dated: February 2, 2007

FEDERAL EXPRESS CORPORATION

By: /s/ Frederick L. Douglas

Frederick Douglas
FEDERAL EXPRESS CORPORATION
3620 Hacks Cross Road
Building B, 3rd Floor
Memphis, Tennessee, 38125

Gilmore F. Diekmann, Jr.
Seyfarth Shaw LLP
560 Mission Street, Suite 3100
San Francisco, California 94105

Evelyn L. Becker
O'Melveny & Myers LLP
1625 Eye Street, NW
Washington, DC 20006

Attorneys for Defendant Federal Express Corporation

Doc. No. 632070

1 Frederick Douglas
(Admitted Pro Hac Vice)
2 FEDERAL EXPRESS CORPORATION
3620 Hacks Cross Road
3 Building B, 3rd Floor
Memphis, Tennessee, 38125
4 Telephone: (901) 434-8519
Facsimile: (901) 434-9271
5

6 SEYFARTH SHAW LLP
Gilmore F. Diekmann, Jr. (SBN: 050400)
Francis J. Ortman, III (SBN: 213202)
7 560 Mission Street, Suite 3100
San Francisco, California 94105
8 Telephone: (415) 397-2823
Facsimile: (415) 397-8549
9

10 Evelyn L. Becker (CA SBN: 170903)
O'MELVENY & MYERS LLP
1625 Eye Street, NW
11 Washington, DC 20006
Telephone: (202) 383-5300
12 Facsimile: (202) 383-5414

13 Attorneys for Defendant
FEDERAL EXPRESS CORPORATION
14

15 **UNITED STATES DISTRICT COURT**

16 **NORTHERN DISTRICT OF CALIFORNIA**

17 DERRICK SATCHELL, KALINI
BOYKIN, VALERIE BROWN, RICK
18 GONZALES, CYNTHIA GUERRERO,
RACHEL HUTCHINS, DANIEL
19 SHERMAN, KELVIN SMITH, SR., and
KEN STEVENSON, on behalf of
20 themselves and all others similarly situated,

21 Plaintiff,

22 v.

23 FEDEX EXPRESS, a Delaware
corporation,
24

25 Defendant.
26
27
28

Case Nos. C03-02659 SI; C03-02878 SI

**DECLARATION OF EVELYN L. BECKER IN
SUPPORT OF DEFENDANT'S OPPOSITION
TO PLAINTIFFS' MOTION IN LIMINE NO. 4
TO PRECLUDE DEFENDANT'S EXPERT
MARY BAKER'S ANALYSIS AND OPINION
CONTROLLING FOR THE RACE OF
MANAGERS**

DATE: March 2, 2007

TIME: 9:00 a.m.

CTRM: The Hon. Susan Illston

1 I, Evelyn L. Becker, hereby declare as follows:

2
3 1. I am a member in good standing of the State Bar of California. I am a partner in
4 the law firm of O'Melveny & Myers LLP, and one of the counsel for the Defendant Federal
5 Express Corporation, in this case. I make this declaration in support of Defendant's Opposition to
6 Plaintiffs' Motion In Limine No. 4 to Preclude Defendant's Expert Mary Baker's Analysis and
7 Opinion Controlling for the Race of Managers.

8 2. Attached to this Declaration as Exhibit A are relevant portions of the Expert
9 Report of Anthony G. Greenwald, dated September 3, 2006.

10 3. Attached to this Declaration as Exhibit B are relevant portions of the Deposition of
11 Anthony G. Greenwald, Ph.D., dated September 22, 2006 and October 27, 2006.

12 4. Attached to this Declaration as Exhibit C are relevant portions of the Deposition of
13 William T. Bielby, Ph.D., dated November 13, 2006.

14 5. Attached to this Declaration as Exhibit D are relevant portions of the Expert
15 Report of William T. Bielby, Ph.D., dated June 30, 2006.

16 6. Attached to this Declaration as Exhibit E are relevant portions of the December 20,
17 2006 Report of Mary Dunn Baker, Ph.D., *Satchell, et. al. v. FedEx Express*. Additional relevant
18 portions of this Report are attached to the Declaration of James M. Finberg, Docket No. 619.

19 7. Attached to this Declaration as Exhibit F is a copy of K. Kraiger and J.K. Ford, "A
20 Meta-Analysis of Race Effects in Performance Ratings," *Journal of Applied Psychology*, Vol. 70,
21 1985, p. 56-65.

22 I declare under penalty of perjury under the laws of the United States that the foregoing is
23 true and correct.

24 Executed this 2nd day of February, 2007.

25
26 /s/ Evelyn L. Becker
Evelyn L. Becker
27
28

EXHIBIT A

EXPERT REPORT OF ANTHONY G. GREENWALD

Derrick Satchell et al. v. FedEx Express

Anthony G. Greenwald, Ph.D.

A. G. Greenwald

September 3, 2006

TABLE OF CONTENTS

A. Assignment	2
B. Qualifications	2
C. Material Reviewed	4
D. Summary of Conclusions Regarding the Baker Report	5
E. Summary of Conclusions Regarding the Tetlock Report	6
F. Analysis of the Baker Report	8
G. Analysis of the Tetlock Report	21
H. References	37

ATTACHMENTS

1. Greenwald CV
2. Consulting fee schedule
3. Greenwald & Krieger article
(in press in *California Law Review*, August, 2006)
"Scientific Foundations of Implicit Bias"
4. Previous testimony

Implicit Bias: Scientific Foundations

Anthony G. Greenwald†
Linda Hamilton Krieger††

The assumption that human behavior is largely under conscious control has taken a theoretical battering in recent years. Although this assault in some ways resembles the previous century's Freudian revolution, there are important differences between the two. Freud's views of unconscious mechanisms were embedded in a theory that never achieved conclusive support among scientists, despite many empirical theory-testing efforts in the middle third of the twentieth century.¹ Consequently, most psychologists have abandoned Freud's psychoanalytic theory of unconscious mental processes.

Theoretical conceptions of conscious control over human behavior were strongly re-established in the last third of the twentieth century, but the dominance of such views has been crumbling during the past two decades. Unlike the Freudian revolution, however, the new science of unconscious mental process is not the product of a single brilliant theoretical mind. Rather, it is being constructed from an evolving, accumulating body of reproducible research findings.²

Copyright © 2006 California Law Review, Inc. California Law Review, Inc. (CLR) is a California nonprofit corporation. CLR and the authors are solely responsible for the content of their publications.

† Professor of Psychology, University of Washington.

†† Professor of Law, University of California, Berkeley School of Law (Boalt Hall). Thanks to Jennifer Eberhardt, Jerry Kang, Tom Newkirk, and Jeff Rachlinski for comments on preliminary versions of this article, and to Ian Ayres, Gary Blasi, Jack Dovidio, John Jost, and Mahzarin Banaji for useful discussions that preceded the writing.

1. See Matthew Hugh Erdelyi, *Psychoanalysis: Freud's Cognitive Psychology* (1985); Matthew Hugh Erdelyi & Benjamin Goldberg, *Let's Not Sweep Repression Under the Rug: Toward a Cognitive Psychology of Repression*, in *Functional Disorders of Memory* 355 (John F. Kihlstrom & Frederick J. Evans eds., 1979); Anthony G. Greenwald, *New Look 3: Unconscious Cognition Reclaimed*, 47 *Am. Psychol.* 766 (1992); John F. Kihlstrom, *The Psychological Unconscious*, in *Handbook of Personality: Theory and Research* 445 (Lawrence A. Pervin ed., 1990); Howard Shevlin & Scott Dickman, *The Psychological Unconscious: A Necessary Assumption for All Psychological Theory?*, 35 *Am. Psychol.* 421 (1980).

2. The early stages of this modern revolution are reviewed by Greenwald, *supra* note 1. Nisbett and Wilson's exposé of the inadequacies of introspective explanations of behavior was a noticeable starting point of the modern revolution, leading to widespread understanding that the self-report measures of conscious mental process that were widely used in psychological research were highly suspect. See Richard E. Nisbett & Timothy DeCamp Wilson, *Telling More Than We Can Know: Verbal Reports on Mental Processes*, 84 *Psychol. Rev.* 231 (1977). Wegner's and Bargh's more recent works reveal the frequency with which seemingly ordinary voluntary actions are controlled in ways that evade conscious scrutiny, further undermining the idea that a conscious mind is the effective governor of most human behavior. See generally Daniel M. Wegner, *The Illusion of Conscious Will*

This Article introduces *implicit bias*—an aspect of the new science of unconscious mental processes that has substantial bearing on discrimination law. Theories of implicit bias contrast with the “naïve” psychological conception of social behavior,³ which views human actors as being guided solely by their explicit beliefs and their conscious intentions to act. A belief is *explicit* if it is consciously endorsed. An intention to act is *conscious* if the actor is aware of taking an action for a particular reason. Of course, actors may dissemble and deny they are taking an action for a particular reason, so conscious intentions based on explicit beliefs may be hard to verify. But a deceitful actor is nevertheless *capable* of asserting the belief or identifying the intention that provides the basis for action, even when *unwilling* to do so.⁴ In contrast, the science of implicit cognition suggests that actors do not always have conscious, intentional control over the processes of social perception, impression formation, and judgment that motivate their actions.

I

Implicit Cognition

Many mental processes function implicitly, or outside conscious attentional focus.⁵ These processes include implicit memory,⁶ implicit per-

(2002); John A. Bargh et al., *The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals*, 81 J. Personality & Soc. Psychol. 1014 (2001).

3. “Naïve psychology” refers to laypersons’ intuitions about determinants and consequences of human thought and behavior, especially their own. Modern treatments were largely inspired by Fritz Heider’s book, *The Psychology of Interpersonal Relations*, which initiated systematic investigation of how laypersons’ intuitions differ from scientific understanding. Fritz Heider, *The Psychology of Interpersonal Relations* (1958).

4. Methodological investigations by social psychologists in the 1960s revealed social influences operating within research and interview settings that would lead people to describe their explicit beliefs inaccurately in experimental studies. See Martin T. Orne, *On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications*, 17 Am. Psychol. 776 (1962); Milton J. Rosenberg, *The Conditions and Consequences of Evaluation Apprehension*, in *Artifact in Behavioral Research* 279 (Robert Rosenthal & Ralph L. Rosnow eds., 1969); Stephen J. Weber & Thomas D. Cook, *Subject Effects in Laboratory Research: An Examination of Subject Roles, Demand Characteristics, and Valid Inference*, 77 Psychol. Bull. 273 (1972). Work inspired by Festinger’s cognitive dissonance theory initiated modern interest in understanding people’s inability to identify the causes of their own thought and behavior. See Leon Festinger, *A Theory of Cognitive Dissonance* (1957). Nisbett and Wilson’s article summarizes the humbling implications of the ensuing two decades of research. See Nisbett & Wilson, *supra* note 2.

5. For an overview of implicit social cognition, which encompasses the phenomena of implicit attitudes, stereotypes, self-esteem, and self-concept, see Anthony G. Greenwald & Mahzarin R. Banaji, *Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes*, 102 Psychol. Rev. 4 (1995); see also Anthony G. Greenwald et al., *A Unified Theory of Implicit Attitudes, Stereotypes, Self-Esteem, and Self-Concept*, 109 Psychol. Rev. 3 (2002).

6. See Daniel L. Schacter, *Implicit Memory: History and Current Status*, 13 J. Experimental Psychol.: Learning, Memory, & Cognition 501 (1987); Larry L. Jacoby & Mark Dallas, *On the Relationship Between Autobiographical Memory and Perceptual Learning*, 110 J. Experimental Psychol.: General 306 (1981).

whether individual differences in implicit attitudes or stereotypes measured by the IAT correlate with (i.e., are predictive of) individual differences in behavior.

A recent meta-analytic review by Poehlman, Uhlmann, Greenwald, and Banaji identified and summarized sixty-one such studies.³⁰ This review's method appraised the value of IAT measures by assessing the relevant body of research in the aggregate, rather than as isolated research findings. To do this, the researchers averaged the available correlational measures of relations between IAT measures and behaviors of interest within groups of studies that tested related hypotheses, as well as over the entire group of eighty-six independent findings from the sixty-one studies. For comparison, parallel analyses examined the aggregated correlations of the same behavioral measures with self-report (explicit) measures, which most of the studies had also obtained.

Both the implicit (IAT) and the parallel explicit measures displayed *predictive validity*, meaning that both types of measures, on average, were significantly correlated with measures of behavior, as expected. To be clear, this does not mean that statistically significant correlations were found in all studies, but that averages of the correlational results of the collected similar tests clearly showed the expected relationships. Predictive validity was greater (meaning that the average correlation was larger) for self-report (explicit) measures than for IAT measures. However, within the critical group of studies that focused on prejudicial attitudes and stereotypes—in other words, within the studies of implicit bias—*predictive validity was significantly greater for the IAT measures*.

Importantly, implicit measures of bias have relatively greater predictive validity than explicit measures in situations that are socially sensitive, like racial interactions, where impression-management processes might inhibit people from expressing negative attitudes or unattractive stereotypes. Additionally, implicit measures have relatively greater validity in predicting spontaneous behaviors such as eye contact, seating distance, and other such actions that communicate social warmth or discomfort.

VI

How Pervasive Is Implicit Bias?

Since 1998, IAT measures of implicit attitudes have been available on the Internet for self-administered demonstration use.³¹ These Web-accessible demonstrations, which allow users to interactively experience

30. See T. Andrew Poehlman et al., *Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity* (2005) (unpublished manuscript, on file with authors).

31. Interactive demonstrations of more than a dozen versions of the IAT are available at <https://implicit.harvard.edu>.

the IAT, have accumulated sufficient data to allow researchers to draw conclusions about the pervasiveness of implicit and explicit biases.³²

Table 1 displays results for a dozen data sets, comparing the degree of favoritism toward advantaged versus disadvantaged groups revealed by implicit versus explicit measures. Two differences between the implicit and explicit measures are readily apparent in these data. First, the explicit measures generally show much greater evidence for attitudinal impartiality or neutrality. Averaged across the dozen topics, 42% of respondents expressed exact or near-exact neutrality on explicit measures. On the IAT measures, however, only 18% of respondents demonstrated sufficiently small implicit bias to be judged implicitly neutral. Second, the IAT measures consistently revealed greater bias in favor of the relatively advantaged group (averaging almost three-quarters of respondents across all the topics) than did the explicit measures (for which an average of slightly over one-third of respondents showed bias favoring advantaged groups).

Table 1 also shows a bias index, computed as the percentage of respondents showing favorability to the advantaged group minus the percentage showing favorability to the disadvantaged group. Whereas this index averaged only 20% for explicit measures, it averaged the very large value of 64% for IAT measures. The broad generalization justified by the data in Table 1 is that implicit attitude measures reveal far more bias favoring advantaged groups than do explicit measures.

It is important to note that these data came from voluntary visitors to the IAT website—a *self-selected* sample, which is different from a *representative* sample that can be obtained by selecting and recruiting respondents randomly from a defined population. As a result, the data in Table 1 cannot be interpreted as representing the attitude distribution of some specific population of interest, such as adult residents of the United States. Even so, the greater favoritism to advantaged groups found in IAT measures than in explicit measures would almost certainly be found with representative samples. Strong evidence for this assertion comes from examination of the Race IAT data for the wide-ranging demographic subgroups shown in Table 2.

Table 2 shows that, with one notable exception, the percentage of respondents who display implicit race bias varies relatively little across groups categorized by varied age, sex, and educational attainment. African Americans constitute the *only* subgroup of respondents who do not show substantial implicit pro-EA race bias on the Race IAT. Approximately equal percentages of African Americans displayed implicit bias in the pro-AA and pro-EA directions. Significantly, among African Americans, IAT

32. These demonstration tests were not set up to conduct research but were nevertheless obliged to record data to enable computations of results that were reported to web visitors. The accumulated data provided by the site's many visitors has proved to be a remarkably rich archive.

results showed considerably greater favoritism to the dominant European American group than did the results from self-report measures, which showed very strong favoritism toward African Americans. The results shown in Table 2 strongly suggest that *any* non-African American subgroup of the United States population will reveal high proportions of persons showing statistically noticeable implicit race bias in favor of EA relative to AA.

Subcategories	N	Self-report (Explicit)				IAT (Implicit)			
		Percent favoring				Percent favoring			
		AA	neither	EA	Index	AA	neither	EA	Index
Education Level									
thru high school grad	3869	9.9%	57.9%	32.2%	22%	9.8%	26.2%	64.0%	54%
at least some college	13028	11.3%	54.1%	34.6%	23%	10.2%	23.2%	66.6%	56%
at least some grad school	3829	12.5%	53.5%	34.0%	21%	12.4%	24.8%	62.9%	50%
Race and Ethnicity									
Black (incl. multiracial)	2277	58.9%	36.2%	4.8%	-54%	34.1%	33.6%	32.4%	-2%
Hispanic (not Black)	1204	15.0%	59.7%	25.3%	10%	10.2%	29.2%	60.5%	50%
Asian & Pacific Islander	1080	9.6%	57.5%	32.9%	23%	7.7%	24.8%	67.5%	60%
White	14805	3.4%	56.0%	40.7%	37%	6.8%	21.7%	71.5%	65%
Age									
under 25	13823	9.7%	55.7%	34.5%	25%	9.4%	23.7%	66.9%	58%
25-44	5403	14.9%	53.9%	31.2%	16%	12.8%	24.4%	62.8%	50%
45 and older	1743	12.3%	47.1%	40.6%	28%	12.6%	25.6%	61.8%	49%
Sex									
Female	13060	12.3%	57.8%	29.8%	17%	11.4%	25.2%	63.4%	52%
Male	7971	9.6%	49.4%	41.0%	31%	9.2%	22.2%	68.6%	59%
Political Ideology									
Conservative	3053	4.8%	44.0%	51.2%	46%	6.5%	19.9%	73.6%	67%
Middle	10612	11.0%	54.0%	35.0%	24%	10.3%	23.8%	65.9%	56%
Liberal	6427	14.8%	59.9%	25.3%	11%	12.9%	26.0%	61.1%	48%

The finding of high levels of the bias index for all demographic subgroups other than Black (i.e., African American) indicates the pervasiveness of pro-EA bias. Even though the bias index was lower in groups of Hispanics and political liberals than in other groups, it was still quite high among those groups.

The finding of high levels of the bias index for all demographic subgroups other than Black (i.e., African American) indicates the pervasiveness of pro-EA bias. Even though the bias index was lower in groups of Hispanics and political liberals than in other groups, it was still quite high among those groups.

EXHIBIT B

WORKING COPY

Page 1

Page 3

Page

EXAMINATION BY MR. DIEKMANN 5

EXHIBITS MARKED FOR IDENTIFICATION

Exhibit No.	Description	Page
1	Expert Report of Anthony G. Greenwald, Derrick Satchell et al. v. FedEx Express	5
2	E-mail correspondence between Anthony Greenwald and Mary Jackman	19
3	Correspondence between Anthony Greenwald, Ph.D., and Journal of Personality and Social Psychology	34

CERTIFIED QUESTIONS

Page 2

1 THE VIDEOGRAPHER: Ladies and gentlemen, we
2 are on video record on September 22nd, 2006. The time
3 is 9:02. I am Jake Krohn, a notary public for the
4 County of Los Angeles, representing Tooker & Antz, 350
5 Sansome Street, Suite 700, San Francisco, California
6 94104, (415) 392-0650. The court reporter is Cheryl
7 Wilder, also representing Tooker & Antz.

8 This is the beginning of Tape 1, Volume 1 in
9 the case of Satchell, et al. vs. FedEx Express in the
10 United States District Court, Northern District of
11 California, Case No. C03-2659 SI and C03-2878 SI for the
12 deposition of Dr. Anthony Greenwald. We are located at
13 560 Mission Street, San Francisco, California. The
14 deposition was noticed by attorneys for defendant and
15 the videotape is produced by defendant.

16 Counsel, would you please identify yourselves
17 and your clients?

18 MR. FINBERG: Jim Finberg for the plaintiff
19 classes.

20 MR. DIEKMANN: And Gil Diekmann, defendant --
21 for defendant Federal Express.

22 THE VIDEOGRAPHER: Okay. The reporter may
23 swear in the witness.

24 (Witness sworn.)

Page 4

VIDEOTAPED DEPOSITION OF ANTHONY GREENWALD, Ph.D. - 9/22/2006

1 group and that they have more contact with this group
2 than do other segments of the population. And I suspect
3 that those African-Americans who show white preference
4 are probably from a portion of the African-American
5 population that has more contact with whites, has more
6 admired whites in their environment than the portion of
7 the black or African-American population that shows
8 implicit preference for black.

9 Q. Are there any studies that bear that out or is
10 that just an assumption on your part?

11 A. Yes. There are some studies that bear this
12 out, but they are not -- the data on this haven't gone
13 as far as one might desire. So yes, there is evidence
14 for this, but one would desire more evidence.

15 Q. I take it that all segments includes the jury
16 who is going to decide this case?

17 A. All segments --

18 Q. Of the population.

19 A. -- have implicit -- well, I would say unless
20 the jury is 100 percent black, one would expect it to
21 have implicit bias.

22 MR. FINBERG: We'll stipulate to that. We
23 oblige you.

24 THE WITNESS: Yes. If you wanted to select a
25 jury that you could be sure was lacking in net implicit

Page 89

1 are found in the larger population.

2 Q. BY MR. DIEKMANN: This whole area of what you
3 call implicit bias is basically the area in which you
4 specialize; is that correct?

5 A. I -- it is certainly a large part of what I
6 specialize in.

7 Q. And you in fact invented a purported test of
8 implicit bias?

9 A. Well, I don't describe it that way, as
10 purported. But yes, I did invent a test of implicit
11 assoc -- that is used as a measure of implicit bias and
12 which I am willing to refer to that way.

13 Q. Okay. And that is the implicit association
14 test?

15 A. It's some variation of the implicit
16 association test.

17 Q. And you've written a large number of articles,
18 both presenting your -- the implicit bias theory and
19 defending it?

20 A. That's not quite accurate, because there's --
21 I'm not sure I have an implicit bias theory. So the
22 articles I have written have mostly been empirical in
23 focus, not all, some have been theoretical, but there is
24 no such thing as implicit bias theory, there is a
25 method. It's really -- the articles are mostly focused

Page 91

1 bias one way or the other towards white and black, our
2 data shows the best you could do is select a 100 percent
3 black jury.

4 Q. BY MR. DIEKMANN: Do you think the fairest --
5 the group who would most fairly analyze a racial issue
6 would be the African-American group?

7 A. I wouldn't go that far, because I think people
8 are capable of being fair independent of their implicit
9 biases. I would certainly hope so.

10 Q. You wouldn't expect the African-American
11 community to have biases or prejudices against
12 African-Americans, would you?

13 A. Well, we've already covered this I think a few
14 times and the answer is yes, I would expect them to have
15 biases against African-Americans.

16 Q. Do you believe that the court system generally
17 is biased against African-Americans?

18 MR. FINBERG: If you have an opinion.

19 THE WITNESS: I -- the only way I can comment
20 on this is to generalize from what I know about the
21 larger population, but in addition, I have done some
22 presentations to groups of judges and who have taken
23 IATs, and my expectation is that judges, juries, and
24 attorneys are probably reflective of the larger
25 population and would have the same kinds of biases that

Page 90

1 on a method.

2 Q. Well, I guess that gets us back to what we
3 were talking about earlier, is that a -- isn't there a
4 theory about what tests like the implicit association
5 test in fact measures?

6 A. Yes. There is undoubtedly theory there and
7 that is something that I have discussed in a few places,
8 including the Greenwald and Krieger article that is an
9 attachment to my report. And the theory is not a
10 complex one. The theory is the theory that it measures
11 associative knowledge.

12 Q. You know, at least -- your answer confuses me,
13 because I've read some articles that say basically what
14 it does measure is not bias at all, but associative
15 knowledge. And that seems to be the side of the debate
16 in this area that is on the other side of the theory
17 that I, at least, understood you to espouse was what is
18 being measured is essentially a predisposition to engage
19 in discriminatory behavior.

20 A. The last part of this predisposition to engage
21 in discriminatory behavior is what I regard as an
22 empirical question. And so we, what we want to know is
23 what is the connection between what this test measures
24 and behavior. It's not quite -- that relationship is
25 itself not part of the theory underlying the test. So

Page 92

23 (Pages 89 to 92)

VIDEOTAPED DEPOSITION OF ANTHONY GREENWALD, Ph.D. - VOL. II - 10/27/2006

1 UNITED STATES DISTRICT COURT
2 NORTHERN DISTRICT OF CALIFORNIA

3 —o0o—

4 DERRICK SATCHELL, KALINI BOYKIN,
5 VALERIE BROWN, RICK GONZALES,
6 CYNTHIA GUERRERO, RACHEL HUTCHINS,
7 KELVIN SMITH, SR., and KEN
8 STEVENSON, on behalf of
9 themselves and all others
10 similarly situated
11 Plaintiffs,

12 vs. No. C 03-2659 SI;
13 C 03-2878 SI

14 FEDEX EXPRESS, a Delaware
15 corporation,

16 Defendants.

17 VIDEOTAPED DEPOSITION OF ANTHONY G. GREENWALD, Ph.D.
18 Volume II
19 Page 265 - 482
20 Friday, October 27, 2006

21 REPORTED BY:

22 NANCY ANN WYNKOOP, CSR #9275

23 TOOKER & ANTZ
24 COURT REPORTING & VIDEO SERVICES
25 350 SANSOME STREET, SUITE 700
SAN FRANCISCO, CALIFORNIA 94104
(415) 392-0650

Page 265

1 BE IT REMEMBERED THAT, pursuant to Notice of
2 Taking Deposition, and on Friday, October 27, 2006,
3 commencing at the hour of 9:00 a.m., at 560 Mission
4 Street, San Francisco, California, before me, NANCY
5 ANN WYNKOOP, duly authorized to administer oaths
6 pursuant to Section 2093 (b) of the California Code of
7 Civil Procedure, personally appeared

8 ANTHONY G. GREENWALD, Ph.D.,
9 called as a witness by the Defendants, and
10 the said witness, having previously stated he would
11 testify to the truth, the whole truth, and nothing but
12 the truth, was thereupon examined and testified as
13 hereinafter set forth.

14 LAW OFFICES OF LIEFF, CABRASER, HEIMANN &
15 BERSTEIN, LLP, 275 Battery Street, 30th Floor, San
16 Francisco, California 94111, represented by JAMES M.
17 FINBERG, ATTORNEY AT LAW, appeared as counsel on
18 behalf of the Plaintiffs.

19 LAW OFFICES OF SEYFARTH SHAW, LLP, 560
20 Mission Street, San Francisco, California 94105,
21 represented by GILMORE DIEKMANN, ATTORNEY AT LAW,
22 appeared as counsel on behalf of the Defendants.

23 Also present: Jake Krohn, videographer.

24 —o0o—

Page 267

1 I-N-D-E-X

2 —o0o—

3
4 DEPOSITION OF ANTHONY G. GREENWALD, Ph.D.

5
6 EXAMINATION BY: PAGE
7 MR. DIEKMANN (Resumed) 268

8 —o0o—

9 DEFENDANT'S EXHIBITS:

10 4 Undated letter to Dr. Banaji from Harris
11 Cooper and attached reviews, 13 pages,
12 inclusive 358

13 —o0o—

Page 266

1 THE VIDEOGRAPHER: Ladies and gentlemen, we
2 are on video record on 10/27/2006, and the time is
3 9:00 o'clock. I'm Jake Krohn, a Notary Public for the
4 County of Los Angeles, representing Tooker & Antz, 350
5 Sansome Street, Suite 700, San Francisco, California
6 94104, (415) 392-0650. The court reporter is Nancy
7 Wynkoop, also representing Tooker & Antz.

8 This is the beginning of Tape One, Volume
9 II, in the case of Satchell, et al. versus FedEx
10 Express, for the deposition of Dr. Anthony Greenwald.
11 All aspects as indicated on Tape One, volume I, remain
12 the same. Counsel may proceed.

13 —o0o—

14 ANTHONY G. GREENWALD, Ph.D.,
15 having previously stated he would testify to the
16 truth, the whole truth and nothing but the truth,
17 testified as follows:

18 EXAMINATION BY MR. DIEKMANN (Resumed)

19 MR. DIEKMANN: Q. Good morning,
20 Dr. Greenwald. I just want to simply remind you that
21 you're still under oath.

22 A. Yes.

23 Q. The fact that you haven't been resworn
24 doesn't mean that the oath you took at the beginning
25 of the first day doesn't still apply.

Page 268

1 (Pages 265 to 268)

VIDEOTAPED DEPOSITION OF ANTHONY GREENWALD, Ph.D. - VOL. II - 10/27/2006

<p>1 housing discrimination have declined since the laws 2 were enacted?</p> <p>3 MR. FINBERG: Objection.</p> <p>4 THE WITNESS: I was trying to answer that, 5 and you told me that what I was suggesting was 6 irrelevant.</p> <p>7 MR. DIEKMANN: Q. No, I'm -- you want -- 8 you gave me some studies that all exist -- that all 9 occurred after the laws were enacted.</p> <p>10 A. But they could be declining during that 11 period of during the period of those studies.</p> <p>12 Q. Doctor, I'm testing your statement that 13 legal restrictions have a minimal or minuscule or very 14 small impact on people's behavior.</p> <p>15 MR. FINBERG: Well, and he's answered it 16 directly responsive to that.</p> <p>17 MR. DIEKMANN: He hasn't.</p> <p>18 Q. So, what I want to know is whether you have 19 any knowledge of the degree to which discrimination in 20 housing declined after the laws were enacted to 21 prohibit it.</p> <p>22 MR. FINBERG: Objection, asked and 23 answered.</p> <p>24 THE WITNESS: Yeah, they've declined 25 somewhat, as evidenced by the fact that during the</p> <p style="text-align: right;">Page 457</p>	<p>1 discrimination, not residential patterns.</p> <p>2 A. Well, you said that you were aware that 3 discrimination had declined, and I was asking you what 4 the evidence was, given the fact that residential 5 segregation is approximately the same as it was 50 6 years ago.</p> <p>7 Q. What I understand and what I believe and 8 what I know is irrelevant. I'm asking whether you are 9 aware of any evidence that indicates that housing 10 discrimination has declined significantly since the 11 laws prohibiting housing discrimination were enacted.</p> <p>12 MR. FINBERG: Objection, asked and 13 answered.</p> <p>14 THE WITNESS: Yeah, I thought I had 15 answered it.</p> <p>16 MR. FINBERG: He said it declined somewhat 17 and it still exists. He wasn't eliminating it. Why 18 don't you move on to another topic, because you're 19 beating this one to death.</p> <p>20 THE WITNESS: I don't think I have anything 21 to add to the answer that I've already given.</p> <p>22 MR. DIEKMANN: Q. Which is, you're aware 23 that it has declined somewhat based on studies all of 24 which were undertaken long after the statute was 25 enacted?</p> <p style="text-align: right;">Page 458</p>
<p>1 period that the law was in effect there have been 2 small declines.</p> <p>3 MR. DIEKMANN: Q. Do you have any idea what 4 the decline has been from the time before the laws 5 were enacted to the time after the laws were enacted?</p> <p>6 A. No, I don't. But I thought I was answering 7 your question nevertheless.</p> <p>8 Q. Do you recognize, based upon the fact that 9 you have been living through the 20th Century, or the 10 last half of the 20th Century, that decline in housing 11 discrimination has been overwhelmingly large since the 12 laws against it have been enacted?</p> <p>13 MR. FINBERG: Objection, assumes facts.</p> <p>14 THE WITNESS: I would like to know the 15 evidence for that, because my assumption is that 16 residential segregation of housing is approximately at 17 the same level it has been for the last half century.</p> <p>18 MR. DIEKMANN: Q. Due to economic 19 circumstances or to intentional discrimination and 20 exclusion?</p> <p>21 MR. FINBERG: Objection, lacks foundation.</p> <p>22 THE WITNESS: You're changing the question. 23 I thought I answered the previous question. Is there 24 another question?</p> <p>25 MR. DIEKMANN: Q. We're talking about</p> <p style="text-align: right;">Page 458</p>	<p>1 A. I am aware that the -- what I said was, the 2 audit studies conducted starting in the '70s showed a 3 level of discrimination in housing in response to a 4 pair of testers that appears to have declined slightly 5 between the '70s and the beginning of the 21st 6 Century. But that itself is not a conclusive answer 7 to the question about discrimination, because there 8 are other forms of discrimination that could occur in 9 a compensatory fashion.</p> <p>10 Q. Page 11, you have this statement: "Only 18% 11 of respondents demonstrated sufficiently small 12 implicit bias to be judged implicitly neutral." And 13 respondents here are respondents to the IAT?</p> <p>14 A. Right.</p> <p>15 Q. So, you're saying that 82% of the population 16 are judged by your theories to be implicitly biased?</p> <p>17 A. One way or the other. Not all in the same 18 direction.</p> <p>19 Q. Table 2 is on Page 14. Directing your 20 attention to the "Race and Ethnicity" section. And 21 does that show that 58.9% of African-Americans on 22 explicit bias testing expressed favoritism toward 23 African-Americans?</p> <p>24 A. Correct.</p> <p>25 Q. And only 4.8% expressed favoritism toward</p> <p style="text-align: right;">Page 460</p>

49 (Pages 457 to 460)

VIDEOTAPED DEPOSITION OF ANTHONY GREENWALD, Ph.D. - VOL. II - 10/27/2006

1 European Americans?
 2 A. Correct.
 3 Q. And does it show that on your IAT scores
 4 34.1% of African-Americans expressed favoritism toward
 5 African-Americans and 33.6% expressed favoritism
 6 toward neither African-Americans nor European
 7 Americans?
 8 A. That's correct.
 9 Q. So, approximately 68% of African-Americans
 10 on the IAT either are neutral or favor
 11 African-Americans?
 12 A. Correct.
 13 Q. Bottom of Page 20 of your article, you
 14 state, "Research has shown that, when a person forms a
 15 new personal connection with a member of a previously
 16 devalued out group, implicit attitudes toward that
 17 group may change dramatically and rapidly." Do you
 18 agree with that?
 19 A. Yes.
 20 Q. Would a new personal connection be working
 21 together with somebody?
 22 MR. FINBERG: Incomplete hypothetical.
 23 THE WITNESS: That's not what I had in
 24 mind.
 25 MR. DIEKMANN: Q. All other forms of new

Page 461

1 A. I wouldn't -- I wouldn't exclude it, but by
 2 no means do I think that it is routinely a personal
 3 connection. But I do think it can be.
 4 Q. Well, there is a big difference between
 5 working down the hall, around the corner and out of
 6 sight, and then working right next to each other on an
 7 open floor, right?
 8 MR. FINBERG: Incomplete hypothetical.
 9 THE WITNESS: I actually haven't had the
 10 experience or done the research that would allow me to
 11 answer that question.
 12 MR. DIEKMANN: Q. Wouldn't the phenomenon
 13 that you report in this sentence -- that is, a close
 14 personal connection with a member of a different
 15 ethnic group -- make a truly integrated workplace one
 16 of the least likely places to find implicit or
 17 explicit biases at work?
 18 A. In connection with this sentence, there is a
 19 footnote, and which amplifies the statement. Can I
 20 read that footnote? Because I think it's relevant
 21 here.
 22 Q. Well, you can do whatever you want. It's
 23 your answer. But, if you're going to read George
 24 Orwell, I'd just as soon not.
 25 A. It's a short footnote.

Page 463

1 personal connections, though, right?
 2 MR. FINBERG: Objection, ambiguous,
 3 compound.
 4 THE WITNESS: I was thinking, as an
 5 example, your white daughter marries an
 6 African-American. That would be a new personal
 7 connection.
 8 MR. DIEKMANN: Q. How about if she just
 9 starts dating an African-American?
 10 A. Well, I have two white daughters. I
 11 wouldn't take that as a personal connection. If --
 12 on the other hand, if they got married or had a child
 13 with and I had a grandchild who was mixed race, I
 14 would consider that a personal connection.
 15 Q. How about if an African-American family
 16 moved next door to you and lived next door to you as
 17 your neighbor for an extended period of time? Would
 18 that be a new personal connection?
 19 A. It could be. I mean, that's the kind of
 20 relationship that has the potential to go one way or
 21 another.
 22 Q. But somehow you will exclude working closely
 23 together with an African-American on the job, five
 24 days a week, 50 weeks a year, not to be a new personal
 25 connection?

Page 462

1 Q. I find the whole subject Orwellian.
 2 A. Olsen, et al., recently reported that an
 3 implicit indicator of expected anti-out-group racial
 4 bias was absent for college students who had
 5 interracial dating experience. So, what I'm reading
 6 there is the evidence that led me to make that
 7 statement that immediately precedes that footnote.
 8 And that's the type of relationship I was referring
 9 to. An interracial dating relationship to me is
 10 different from someone working on the same floor with
 11 you.
 12 Q. Do you think a dating experience gives one
 13 any more of an interpersonal relation with the person
 14 that you're dating than a day-in/day-out close working
 15 relationship?
 16 MR. FINBERG: Incomplete hypothetical.
 17 THE WITNESS: But --
 18 MR. DIEKMANN: Yeah.
 19 MR. FINBERG: Don't fill in the details.
 20 THE WITNESS: So, it's been a long time
 21 since I did dating, and I trust the same may be true
 22 for you. But, yes, I think this is different.
 23 MR. DIEKMANN: Q. I mean, I just don't
 24 necessarily see it, but maybe I have been -- I have
 25 three grown kids, who dated very casually, and it

Page 464

50 (Pages 461 to 464)

EXHIBIT C

VIDEOTAPED DEPOSITION OF WILLIAM BIELBY - VOL. II - 11/13/2006

1 UNITED STATES DISTRICT COURT 2 NORTHERN DISTRICT OF CALIFORNIA 3 --oOo-- 4 5 DERRICK B. SATCHELL, KALINI 6 BOYKIN, VALERIE BROWN, RICK 7 GONZALES, CYNTHIA GUERRERO, 8 RACHEL HUTCHINS, KELVIN SMITH, 9 SR., and KEN STEVENSON, on 10 behalf of themselves and all 11 others similarly situated, 12 Plaintiffs, 13 vs. No. C 03-2659 SI; 14 C 03-2878 SI 15 16 FEDEX EXPRESS, a Delaware 17 corporation 18 Defendant. 19 20 DEPOSITION OF WILLIAM T. BIELBY 21 Volume II, Pages 269 - 476 22 Monday, November 13, 2006 23 24 REPORTED BY: CYNTHIA LEW, RPR, CSR No. 11999 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000	1 BE IT REMEMBERED that, pursuant to Notice 2 of Taking Deposition, on Monday, November 13, 2006, 3 commencing at the hour of 8:58 o'clock a.m. thereof, 4 at the Law Offices of Seyfarth Shaw, 560 Mission 5 Street, Suite 3100, San Francisco, California 94105 6 before me, CYNTHIA LEW, duly authorized to administer 7 oaths pursuant to Section 2093(b) of the California 8 Code of Civil Procedure, personally appeared for the 9 continuation of his deposition 10 WILLIAM T. BIELBY, 11 called as a witness on behalf of the Defendants, and 12 the said witness, having previously been placed under 13 oath, was thereupon examined and testified further as 14 hereinafter set forth. 15 16 APPEARANCES 17 The Law Offices of Schneider & Wallace, 180 18 Montgomery Street, Suite 2000, San Francisco, 19 California 94104, represented by GUY B. WALLACE and 20 NANCY PARK, Attorneys at Law, appeared as counsel on 21 behalf of the Plaintiffs. 22 The Law Offices of Seyfarth Shaw, 560 23 Mission Street, Suite 3100, San Francisco, California 24 94105, represented by GILMORE F. DIEKMANN, JR., 25 Attorney at Law, appeared as counsel on behalf of the Page 269
1 INDEX 2 3 DEPOSITION OF WILLIAM T. BIELBY 4 5 EXAMINATION BY: PAGE 6 MR. DIEKMANN (Resumed) 273 7 8 --oOo-- 9 10 EXHIBITS 11 (For Defendants) 12 IDENTIFICATION DESCRIPTION PAGE 13 4 11-page "Order re Motions to Strike 425 14 Expert Testimony," re: Goshu v. U.S. 15 Bancorp, Dated October 1, 2002 16 5 18-page "Opinion & Order" re: EEOC v. 430 17 Morgan Stanley, dated July 2, 2004 18 19 --oOo-- 20 21 22 23 24 25 Page 270	1 Defendants. 2 Also present: Philip Tetlock and Jake 3 Krohn, videographer. 4 --oOo-- 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 Page 272

1 (Pages 269 to 272)

VIDEOTAPED DEPOSITION OF WILLIAM BIELBY - VOL. II - 11/13/2006

1 me that there's — the sample sizes are large enough
2 that one could do such an analysis. I just don't
3 know how it would turn out.
4 MR. DIEKMANN: Q. Well, what you've seen
5 so far, or at least from the FedEx database, is a
6 pattern similar to the Roth metaanalysis; correct?
7 Neither one of which involve a significant
8 within-subject analysis; right?
9 A. Correct.
10 Q. Another piece of evidence that I want to
11 know whether you know about, did you read Dr. — have
12 you read Dr. Greenwald's report in this case?
13 A. Yes.
14 Q. Did you read the law review article that he
15 attached to the report?
16 A. Not closely. I've seen versions of it over
17 the past few months.
18 Q. Okay. You've seen the table that he
19 includes in there, the big table that, on one side,
20 shows measures of explicit bias and then compares
21 them on the other side to measures of implicit bias
22 for various categories of analysis?
23 A. I have a vague recollection of it, yes.
24 Q. And with respect to African American
25 viewpoints, did you see that an extremely high

Page 445

1 percentage of African Americans express in-group
2 favoritism and preference for African Americans on
3 explicit measures?
4 A. I think it was maybe a third or something
5 like that.
6 Q. You're talking now about implicit, I think.
7 Didn't you see the — well, I think it was almost 70
8 or 80 percent, maybe 60 percent of African Americans
9 expressed in-group favoritism toward African
10 Americans on explicit tests?
11 A. I'm sorry.
12 Q. Explicit measures?
13 A. I misunderstood. You're correct. Yes.
14 Q. And did you see that on the — even based
15 on Dr. Greenwald's implicit-association test Web site
16 that African Americans — the two thirds of African
17 Americans either favor African Americans or are
18 neutral?
19 A. I think that's roughly what I recall, yes.
20 Q. Isn't that a body of — well, do you accept
21 the table as being based upon some recognizable body
22 of social science research?
23 A. Yes.
24 Q. Isn't that some social science research
25 that indicates that African Americans aren't likely

Page 446

1 to be biased against other African Americans overall
2 on a systemic basis?
3 A. I'll go part of the way there. It suggests
4 that African Americans are, in general, less likely
5 to be biased against African Americans than
6 non-African Americans.
7 Q. Well, it also suggests, does it not, that
8 they are — if you follow the explicit-measure
9 studies, they're more likely to be biased in favor of
10 African Americans. Right?
11 A. I believe so, yes.
12 Q. And if you base your analysis on the
13 implicit tests, you would be inclined to conclude
14 that they either were neutral or favored African
15 Americans two to one?
16 A. Well, you could also say neutral or
17 disfavored two to one.
18 Q. True. But I'm talking about the absence of
19 bias here. And so one would say they're at least two
20 to one unlikely to be biased against African
21 Americans. Correct?
22 A. Correct.
23 Q. Do you have any personal belief that
24 African Americans share the same African American
25 stereotypes as whites?

Page 447

1 A. In terms of the content of stereotypes,
2 they are — yeah. There's no reason to believe they
3 would be substantially different.
4 Q. Why not? It would seem to me that exposure
5 alone would result in at least different stereotypes,
6 if not the absence of stereotypes.
7 A. I'm not aware of any research that shows
8 that.
9 Q. What about common sense?
10 A. Well, I mean, there's a body of research
11 that suggests that the way particularly implicit
12 kinds of stereotypes are absorbed is by looking at
13 disparities in the real world and then extrapolating
14 from them. So for example, the stereotype that
15 African Americans lack initiative and don't work hard
16 and so on is a way of making sense of the fact that
17 people — that poor people are disproportionately
18 African American and reinforces — and the objective
19 fact can reinforce the stereotype but have that
20 effect on African Americans as well as whites.
21 So — I mean, this is an empirical question
22 and I would guess there's research out there on it.
23 I just don't know the results of it.
24 Q. Well, I'm not sure it is an empirical
25 question because we're talking about presence of

Page 448

45 (Pages 445 to 448)

EXHIBIT D

1 James M. Finberg (SBN 114850)
 Bill Lann Lee (SBN 108452)
 2 Eve H. Cervantez (SBN 164709)
 Chimène I. Keitner (SBN 226948)
 3 LIEFF, CABRASER, HEIMANN &
 BERNSTEIN, LLP
 4 275 Battery Street, 30th Floor
 San Francisco, CA 94111-3339
 5 Telephone: (415) 956-1000
 Facsimile: (415) 956-1008

6 Todd M. Schneider (SBN 158253)
 7 Guy B. Wallace (SBN 176151)
 Joshua Konecky (SBN 182897)
 8 SCHNEIDER & WALLACE
 180 Montgomery Street, Suite 2000
 9 San Francisco, CA 94104
 Telephone: (415) 421-7100
 10 Facsimile: (415) 421-7105

11 JOHN BURRIS (SBN 69888)
 7677 Oakport Bldg., Suite 1120
 12 Oakland, CA 94612
 Telephone: (510) 839-5210
 13 Facsimile: (510) 839-3882

14
 15 *Class Counsel*

MICHAEL S. DAVIS (SBN 160045)
 345 Hill Street
 San Francisco, CA 94109
 Telephone: (415) 282-4315
 Facsimile: (415) 358-5576

KAY MCKENZIE PARKER (SBN 143140)
 225 Bush Street, 16th Floor
 San Francisco, CA 94104
 Telephone: (415) 227-9622
 Facsimile: (415) 227-4522

WAUKEEN Q. MCCOY (SBN 168228)
 703 Market Street, Suite 1407
 San Francisco, CA 94103
 Telephone: (415) 675-7705
 Facsimile: (415) 675-2530

16 UNITED STATES DISTRICT COURT
 17 NORTHERN DISTRICT OF CALIFORNIA
 18

19 DERRICK SATCHELL, KALINI
 20 BOYKIN, VALERIE BROWN, RICK
 GONZALES, CYNTHIA GUERRERO,
 21 RACHEL HUTCHINS, TYRONE
 MERRITT, KELVIN SMITH, SR., and
 22 KEN STEVENSON, on behalf of
 themselves and all others similarly situated,

23 Plaintiffs,

24 v.

25 FEDEX EXPRESS, a Delaware
 26 corporation,

27 Defendant.
 28

Case No. C 03-2659 SI; C 03-2878 SI

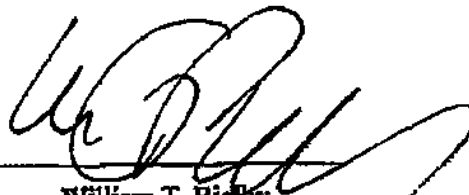
EXPERT REPORT OF
WILLIAM T. BIELBY, PH.D.

EXPERT REPORT OF WILLIAM T. BIELBY

Derrick Satchell, et al. v. FedEx Express

William T. Bielby, Ph.D.

June 2006

A handwritten signature in black ink, appearing to read 'W. T. Bielby', is written over a horizontal line.

William T. Bielby

June 30 2006

**William T. Bielby
16152 First Ln, Box 536
Union Pier, MI 49129-0536
312-497-9184
bielby@sas.upenn.edu**

16. There is ample reason to closely scrutinize any performance review system for potential racial bias. A large body of social science research, including many studies conducted in organizational settings, show that African Americans tend to receive lower performance evaluations than do whites. For example, a 1985 article by Kraiger and Ford performed a meta-analysis of 84 studies with a total of over 20,000 ratees, including 64 field studies and 10 experimental studies of white raters rating a total of 17,159 African American and white ratees.¹⁷ Meta-analysis is a method for quantitatively aggregating results across studies in order to obtain a more precise estimate of the size and reliability of effects than can be obtained from any single study.¹⁸ In addition, a meta-analysis of studies conducted across a variety of settings contributes to establishing the external validity of the research. Kraiger and Ford's meta-analysis found that African Americans, on average, received significantly lower evaluations than did whites. They also found evidence suggesting that white raters give significantly higher ratings to whites than to African Americans; and (2) that African American raters give significantly higher ratings to African Americans than do whites.¹⁹

17. A subsequent meta-analysis published by Martocchio and Whitener in *Human Relations* in 1992 replicates Ford *et al.*'s finding regarding racial differences in

M. O'Rand, "The Cumulative Stratification of the Life Course," p. 188-207 in *Handbook of Aging and the Social Sciences* (4th ed.), edited by R. H. Binstock, L. K. George, et al. Academic Press, 1996.

¹⁷K. Kraiger and J. K. Ford, "A Meta-Analysis of Race Effects in Performance Ratings," *Journal of Applied Psychology*, Vol. 70, 1985, p. 56-65.

¹⁸J. E. Hunter and F. L. Schmidt, *Methods of Meta-analysis: Correcting Error and Bias in Research Findings*, Sage Publications, 1990.

¹⁹The research on in-group favoritism in performance ratings as it applies to minorities is limited because most datasets contain relatively few instances of minority raters completing ratings of both minority and white ratees and minority ratees being rated by both minority and white raters.

EXHIBIT E

REPORT OF MARY DUNN BAKER, Ph.D.

SATCHELL, et al., v. FEDEX EXPRESS

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA**

Case No. C 03 – 2659; C 03-2878 SI

**ERS GROUP
Tallahassee, Florida
December 20, 2006**

- Dionne Dandridge, August 11, 2004;
 - Kathy Michele Davis, July 21, 2004;
 - Jerilyn Hayward, Ph.D., August 20, 2004;
 - Tyrone McQueen, July 15, 2004;
 - Elizabeth F. Powell, July 21, 2004;
 - Danny W. Rhea, July 21, 2004;
 - Gwendolyn Rouzan, August 11, 2004;
 - Robert A. Speroff, July 22-23, 2004;
 - Robin Van Galder, July 15, 2004;
 - Vickie Williamson, August 13, 2004;
 - Valerie Ann Brown, April 26, 2004;
 - Rick Gonzales, May 11, 2004;
 - Cynthia Guerrero, April 21, 2004;
 - Rachel Regina Hutchins, April 23, 2004;
 - Derrick A. Satchell, May 3, 2004;
 - Kelvin James Smith, Sr., Aril 30, 2004;
 - Kenneth Bernard Stevenson, Sr., April 28, 2004; and
 - Tyrone Merritt, February 14, 2006;
- Reports submitted by FedEx's and Plaintiffs' experts; and
 - References to the professional literature cited herein.

III. CONCEPT OF STATISTICAL SIGNIFICANCE

Assuming that employees are similarly situated in terms of relevant factors or characteristics, in a selection process (e.g., promotion or disciplinary action) that is neutral with respect to race or ethnicity, on average, the expected number of protected employees selected will be consistent with their representation in the relevant candidate "pools."⁸ The expected number of protected employees selected is calculated by multiplying the total number of selections by the percent protected among employees in the relevant candidate pools.⁹

⁸ This same concept underlies analyses of continuous variables such as pay and performance ratings.

⁹ In the absence of data regarding the actual candidate pools from which selections were made, in some cases, protected group availability can be estimated using the percent protected among "likely" candidates who are similar to the selected individuals in the relevant respects.

Social scientists and the courts have long recognized that the actual outcome of a neutral selection process may differ from the expected result simply because of random (or chance) variation. In order to determine whether the actual and expected numbers of protected selections are “close enough” to be consistent with the outcome of a neutral selection process, the number of standard deviations of the difference between the actual and expected values (or the probability of chance occurrence) is calculated.

Generally, when the number of standard deviations is less than approximately two (or three) [or the probability of chance occurrence is greater than 5% (or 1%)], social scientists and the courts conclude that the observed difference is consistent with random variation in a neutral system and that the difference is likely to have occurred by chance.¹⁰ Such differences are considered *statistically insignificant* and do not give rise to an inference of discrimination. Therefore, when the number of standard deviations of the difference between the actual and expected number of protected selections is less than two (or three) [or the probability of chance occurrence is greater than 5% (or 1%)], the data fail to provide statistical support for an allegation of discrimination.

When the number of standard deviations is approximately two (or three) or more [or the probability of chance occurrence is 5% (or 1%) or less], typically, social scientists and the courts conclude that the observed difference is “too large” to be consistent with the outcome of a random selection process and that the difference is not likely to have occurred by chance in a neutral setting. Such differences are *statistically significant* and may provide statistical support for allegations of discrimination, assuming that the decision-making process is properly modeled. Therefore, when the outcome is adverse to

¹⁰ For example, see Hazelwood School District v. U.S., 433 U.S. 299, 308n. 14 (1977) and Palmer v. Schultz, 43 FEP Cases, 452 (DC Cir., 1987, pp. 461-465).

the protected group and the number of standard deviations of the difference is two (or three) or more [or the probability of chance occurrence is 5% (or 1%) or less], the data may give rise to an inference of discrimination (again, assuming that the selection process is properly modeled).

Significant differences only indicate that the difference is not likely to have occurred by chance in a neutral setting. Significant results may be reflective of discriminatory decision-making. However, the significant difference, in whole or in part, may be the result of the fact that there are racial or ethnic differences in a factor that influences the variable of interest that happens to be correlated with race or ethnicity and for which the analysis did not account. Therefore, the analyst should consider the impact that omitted variables may have on the outcome of the analysis in the assessment of the meaningfulness of the result.

When analyses uncover statistically significant differences, in addition to considering the likely impact that omitted factors have on the observed result, the analyst should also consider whether the significant difference is important as a practical matter.¹¹ In many situations, especially when the number of observations is large, a difference of a trivial magnitude may be statistically significant. Paetzold and Willborn state,

*"...statistical significance need not correspond to practical significance. Even a trivial disparity can become statistically significant under the right conditions. Thus, a .05 cut-off level need not guarantee that the disparity has any practical significance."*¹²

¹¹ For example, Connolly, Peterson and Connolly state "...at some point, we feel even the most scrupulous adherent to the principles of equality must bow to the realization that degree of inequality is relevant to a finding of liability absent evidence of willful discrimination." [Use of Statistics in Equal Employment Opportunity Litigation, Law Journal Press, 2004.]

¹² The Statistics of Discrimination. Using Statistical Evidence in Discrimination Cases, Clark Boardman Callaghan, 12/1996.

The Federal Judicial Center's Reference Manual on Scientific Evidence also emphasizes that practical significance should be considered.

*"...significant differences are evidence that something besides random error is at work, but they are not evidence that this 'something' is legally or practically important. Statisticians distinguish between 'statistical' and 'practical' significance to make the point. When practical significance is lacking – when the size of a disparity or correlation is negligible – there is no reason to worry about statistical significance."*¹³

Therefore, as these authors caution, in assessing the meaningfulness of a difference, its absolute magnitude should be considered, as well as the number of standard deviations of the difference.

When the number of selections is large and an inferential statistical analysis produces a significant difference between the rates at which protected and non-protected employees are selected, impact ratio analysis is often conducted to determine whether, as a practical matter, the observed difference is "large." Generally, when the selection rate of one group is at least 80% of the other group's rate, the conclusion is that the difference, as a practical matter, is not large.¹⁴

IV. RACIAL AND ETHNIC COMPOSITION OF THE WORKFORCE

Plaintiffs highlight the fact that, in general, the percent minority among employees is higher in the relatively low paying jobs (e.g., Handler, Freight/Material Handler and Checker Sorter) than in relatively high paying positions (e.g., Courier and Operations Manager). Using descriptive statistics, the class implies that the differences

¹³ See "Reference Guide on Statistics" by David H. Kaye and David A. Freedman.

¹⁴ According to the Equal Employment Opportunity Commission, "[a] selection rate for any race, sex or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement authorities as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact."

EXHIBIT F

A Meta-Analysis of Ratee Race Effects in Performance Ratings

Kurt Kraiger
University of Colorado at Denver

J. Kevin Ford
Michigan State University

A recent review of ratee race effects on performance ratings (Landy & Farr, 1980) found conflicting results. For the present research, meta-analytic techniques were used for more substantive conclusions about the existence of ratee race effects and whether the effects were related to rater race or were moderated by situational factors. The five moderators examined included the study setting (laboratory/field), rater training (offered/not offered), type of rating (behavior/trait), rating purpose (administrative/research), and the composition of the workgroup (percentage of blacks in each study). Seventy-four studies with a total sample of 17,159 ratees were located for white raters, whereas 14 studies with 2,428 ratees included data on black raters. The corrected mean correlations between ratee race and ratings for white and black raters were .183 and -.220, respectively, with 95% confidence intervals that excluded zero for both rater groups. Substantial moderating effects were found for study setting and for the saliency of blacks in the sample. Race effects were more likely in field settings when blacks composed a small percentage of the workforce. The appropriate role of laboratory studies and the implications of the results for guiding future research on racial bias are discussed.

A traditional literature review of ratee race effects has been recently completed by Landy and Farr (1980). Many of the studies reviewed showed a significant ratee race effect (deJung & Kaplan, 1962; Farr, O'Leary, & Bartlett, 1971; Hamner, Kim, Baird, & Bigoness, 1974; Landy & Farr, 1973), whereas other studies indicated no effect (Fox & Lefkowitz, 1974; Schmidt & Johnson, 1973). Some studies found effects for rater race (e.g., Campbell, Crooks, Mahoney, & Rock, 1973), whereas others did not (Schmidt & Johnson, 1973). Results of still other studies showed a more complex interaction between rater and ratee race and other variables such as ratee performance level (Bigoness, 1976), age (Toole, Gavin, Murdy, & Sells, 1972), and employment status (Bass & Turner, 1973). Landy

and Farr (1980) concluded that, in general, ratees tend to receive higher ratings from raters of the same race, although situational factors may moderate this effect.

The nature of these diverse findings over various settings, populations, and operationalizations of measures makes any conclusions drawn from a traditional review tenuous at best. The conflicting results call for a more substantive strategy for analyzing the available data on ratee race effects. One such tool is meta-analysis, a method of integrating results from existing studies to reveal patterns of underlying relations and causalities (Hunter, Schmidt, & Jackson, 1982). The results from a meta-analysis not only provide a more definitive appraisal of race effects but allow the researcher to examine the viability of assumptions regarding conditions that may moderate the relationship between ratee race and rating. An examination of the existing rating literature suggests five potential moderators of ratee race effects.

The most recent viewpoint on race effects is that they are much more likely to be found in laboratory than "real world" field settings. Wendelken and Inn (1981) argued that in

This article is based on a paper presented at the 91st Annual Convention of the American Psychological Association, August, 1983, Anaheim, California.

The authors express their thanks to Neal Schmitt, Daniel Ilgen, and two anonymous reviewers for providing comments to earlier drafts of this article.

Requests for reprints should be sent to Kurt Kraiger, Department of Psychology, University of Colorado at Denver, 1100 Fourteenth Street, Denver, Colorado 80202.

typical laboratory studies, ratee race is made salient by artificially constraining ratee performance and by limiting the amount of information available to the rater. In contrast, the real world rater has a larger amount of job-relevant information available that reduces the necessity of using race as a factor in the evaluation. The contention that race is less salient in organizational settings is arguable. Under affirmative action and equal employment pressures, race may become a highly salient and rational consideration in the evaluation of performance in organizations (Mobley, 1982), motivating the rater, either intentionally or unintentionally, to minimize race effects. On the other hand, one could argue effectively that ratee race effects should be less likely in laboratory experiments due to the tendency of college students to provide socially desirable responses (Jones & Sigall, 1971). Consequently, it is unclear whether race effects should be more likely in laboratory or field settings.

A second issue is the impact of rater training on rating effects. Although a number of authors have suggested rater training has or should reduce race effects (e.g., Bigoness, 1976; Hamner, Kim, Baird, & Bigoness, 1974), previous training studies have focused almost exclusively on other systematic rating effects such as halo, leniency, and central tendency. Proponents of this view commonly cite one study by Schmidt and Johnson (1973) for support, but a careful reading of this study reveals that participants were not given any rater training but instead provided ratings after an extensive human relations course as part of a managerial training program. Furthermore, recent empirical evidence indicates that the impact of training on rating effects diminishes over time (Bernardin, 1978; Ivancevich, 1979) and often bears little relationship to the accuracy of ratings (Bernardin & Pence, 1980; Borman, 1979). Thus, although rater training potentially may reduce ratee race effects, there is no empirical evidence that it has, and reasonable cause for pessimism that it could.

A third issue is the extent to which ratee race effects are related to the composition of the workgroup. Although this proposition has not received much attention in the Industrial/Organizational (I/O) literature, it has been

suggested that diminishing the proportion of minority to majority group members should increase the racial saliency of minority members and polarize evaluations (Taylor, 1981). For example, Schmitt and Hill (1977) reported that black female assesses tended to receive lower ratings when their assessment center group was composed principally of white males than when the group was more equally integrated by race and sex. The pervasiveness of the phenomenon across settings is not yet known.

A fourth issue concerns the behavioral specificity of the rating scale. Wherry and Bartlett (1982) have theorized that behaviorally based rating formats will be less prone to rating effects than more global, general formats. Thus, Brugnoli, Campion, and Basen (1979) reported that race effects were minimized when a behaviorally based rating scale was used instead of a global measure. Nevertheless, a comparison of rating formats by Jacobs, Kafry, and Zedeck (1980) found little evidence for the psychometric superiority of behavioral scales over alternative formats. Research on rating formats suggests varying rating formats will do little to reduce ratee race effects.

A final issue often discussed in the rating literature is the impact of the purpose or consequence of ratings on rating effects. As hypothesized by Wherry and Bartlett (1982), ratings obtained under experimental conditions should be more accurate than those obtained under actual job conditions where administrative actions may affect the rater or ratee. Relevant research has shown that ratings are significantly more lenient and exhibit more halo when raters are told the ratings are for administrative rather than research purposes (Sharon & Bartlett, 1969; Taylor & Wherry, 1951; Warmke & Billings, 1979). However, the impact of rating purpose on ratee race effects has not yet been explicitly studied.

In sum, the rating literature has suggested at least five potential moderators of race effects in performance evaluations. Existing empirical evidence for each potential moderator is often limited to sporadic single-setting correlates that may not be generalizable to other situations. The extent and operative nature of each potential moderator

can be investigated either by systematically varying the moderator variable across multiple organizational units or through the cumulation of the results of many studies employing different settings or conditions. Although the former option presents formidable cost and logistical obstacles, the later meta-analytic method is feasible and has already been successfully applied to other domains of I/O psychology (e.g., Schmidt, Hunter, & Pearlman, 1981; Terborg, Lee, Smith, Davis, & Turbin, 1982).

It should be noted at the outset that the present study does not directly isolate the important issue of racial bias in performance evaluations. As Wherry and Bartlett (1982) have stated, subjective ratings are a function of not only ratee performance but also biases in the observation and recall of that performance by the rater. Because individual studies rarely provide information regarding actual performance, a meta-analysis of race effects cannot separate the relative contributions of ratee performance and rater bias to rating differences. Nevertheless, a necessary first step to the investigation of racial bias is to determine whether race is related to evaluations under various conditions.

Thus, meta-analytic techniques are used for more substantive conclusions regarding the existence of ratee race effects and whether the race effects are related to rater race or are moderated by the five situational factors previously discussed.

Method

An attempt was made to locate, summarize, and analyze the results of all published studies and a number of unpublished studies reporting performance ratings of black and white ratees. The search was performed manually using three indexes (*Personnel Literature*, *Personnel Management Abstracts*, and *Psychological Abstracts*) and a systematic review of the *Journal of Applied Psychology* and *Personnel Psychology* from 1966 to 1981. Because misrepresentation of population parameters may result from inadequate sampling procedures (Hunter et al., 1982), a variety of unpublished works was included as well. Sources included solicited responses from researchers active in the areas of test validation and performance evaluation and technical reports readily available to the authors. A total of 34 published and 47 unpublished studies were located for analysis. These totals reflect multiple samples for some reports. Effect sizes were available or determinable from 30 published and 44 unpublished studies, resulting in a total sample size of 74. All studies presented data for white raters, but only 14 studies presented data for black raters. A complete

list of the studies included in the analysis is presented in the Appendix.

Analysis

Different estimates of effect size have been proposed including d (Glass, 1976) and ω^2 (Hays, 1973) where ω^2 is calculated from t . However, both d and t are algebraic transformations of the more generally applied point-biserial correlation, r_{pb} (Hunter et al., 1982). The present meta-analysis cumulated the correlation of ratings and race (arbitrarily coded White = 1, Black = 0) in the computation of the mean effect size (\bar{r}_{pb}) and its variance ($\sigma_{\bar{r}_{pb}}^2$) across studies. An estimate of variance due to sampling error (σ_s^2) and the population variance for effect sizes (σ_e^2) were computed using procedures explained in Hunter et al. (1982). The estimated standard error (σ_e) was used to establish confidence intervals around \bar{r}_{pb} and to test the hypothesis that $\bar{r}_{pb} = 0$ in the population.

Because the size of a point-biserial correlation is affected by the relative proportions of the two groups, effect sizes for individual studies were corrected for differences in subgroup sample sizes (the corrected correlation estimates effect size if subgroup sample sizes were equal). Estimated sampling error was adjusted for added variance due to this correction. Cumulated effect sizes and estimated population variance were also corrected for the average reliability of ratings to estimate true effect size and variance given perfect measurement. The average reported conspect reliability across studies was .70. It should be noted that this value results in a more conservative correction than the .60 value advocated by Hunter & Schmidt (Hunter et al., 1982; King et al., 1980).

Estimated population variance (σ_e^2) represents actual study-to-study variation in effect sizes with estimated variance due to small samples and unequal sample sizes removed. This corrected variance may be trivial in size and likely due to other statistical artifacts or it may be nontrivial and suggest the possible presence of one or more moderators in the data. Two tests of the triviality in corrected variance were performed. The first was a chi-square test of the hypothesis of no variation suggested by Hunter et al. (1982). The second, less formal test was the Schmidt, Hunter, and Pearlman (1982) "bare bones" analysis, which computes the ratio of sampling error variance to true variance to determine whether the observed variance (over 75%) is largely artifactual in nature. The two tests for the triviality of the corrected variance revealed significant chi-squares, $\chi^2(74, N = 17,159) = 221.75, p < .01$, for white raters and $\chi^2(14, N = 2,428) = 76.41, p < .01$, for black raters, and a small sampling error to true variance ratio (less than 40%) for the total sample of studies. These results support the investigation for potential moderators in the data.

For white raters, five potential moderators were investigated. The small number of studies precluded the examination of moderator variables for black raters. Four moderators were coded as dichotomous variables: setting (lab/field), rater training (offered/not offered), type of rating (behavior based/trait), and rating purpose (administrative/research). The coding was completed primarily by the first author. A subset of studies was coded by both authors resulting in a high level (90%) of agreement. The

RATEE RACE EFFECTS

59

Table 1
Mean and Variance of Race Effects in Performance Ratings

	No. studies	Cumulative N	\bar{r}_{pb}	Corrected ^a \bar{r}_{pb}	Estimated d	σ_r^2	σ_p^2	Corrected confidence intervals	Z test ^b
Total sample									7.57**
White raters	74	17,159	.153	.183	.37	.012	.010	.02 < ρ < .35	
Black raters	14	2,428	-.184	-.220	-.45	.029	.032	-.41 < ρ < -.03	
White raters— moderator analyses									
Research setting									2.41*
Laboratory	10	1,010	.031	.037	.07	.021	.002	-.20 < ρ < .27	
Field	64	16,149	.160	.192	.39	.011	.008	.03 < ρ < .35	
Rater training ^c									1.32
Offered	16	5,055	.173	.207	.42	.013	.012		
Not offered	39	10,088	.158	.189	.38	.008	.005		
Rating scales ^c									1.12
Behavioral	32	8,692	.178	.213	.43	.011	.009		
Trait	23	6,451	.142	.170	.34	.007	.003		
Purpose ^c									1.19
Administrative	18	6,955	.166	.199	.41	.008	.006		
Research	37	8,259	.159	.191	.39	.010	.006		

^a Correction for attenuation, average interrater reliability = .70.

^b Z test is for subgroups immediately below test statistic.

^c Variation by subgroup for field studies only.

* $p < .05$. ** $p < .001$.

presence of these moderators was examined by classifying the studies into relevant subsamples and recomputing subsample r_{pb} s and confidence intervals. Differences in subgroup effect sizes were tested for significance by a procedure adapted from Rosenthal and Rubin (1982). The procedure requires the definition of the subgroup comparison by a set of contrast weights such that the sum of the contrasts over all studies equals zero. The procedure also requires the computation of estimated variance of individual study effect sizes. Let L equal the sum of the product of contrast weights times effect sizes over all studies and M equal the sum of the contrasts over studies squared divided by the estimated effect size variance. Rosenthal and Rubin (1982) have shown that the quotient L/M is distributed as a standard normal deviate, Z .

The fifth moderator, racial composition of the workgroup, was investigated by coding the percentage of blacks rated in each study and correlating this value with the uncorrected effect size across all studies (i.e., the effect size before the correction for differences in subgroup sample size). A significant correlation across studies indicates the percentage of blacks rated moderates the degree of relationship between race and performance rating (Arnold, 1982). If significant, the correlation can be corrected for measurement error by the formula provided by Hunter et al. (1982).

Results

Effects for White and Black Raters

The major results of the meta-analysis are presented in Table 1. The table shows cumulated sample sizes, effect sizes, variance

estimates, and confidence intervals for the performance ratings of black and white raters. Results of moderator analyses for white raters are also presented in Table 1.

The best estimate of the population effect size (ρ_{pb}) is the mean point-biserial correlation corrected for unreliability in ratings. For white raters, this estimate was .183 and was based on a sample of 74 studies and 17,159 ratees. The 95% confidence interval about this estimate excluded zero (.02 < ρ < .35), indicating that white raters assigned significantly higher ratings to white ratees than to black ratees. Transforming the corrected \bar{r}_{pb} to Glass's (1976) d statistic (.37) suggests an alternate but compatible interpretation: The average white ratee received a more favorable evaluation from white raters than 64% of the black ratees. The estimated population variance of effect sizes for white raters ($\sigma_p^2 = .010$) was small but greater than zero and suggests the presence of one or more moderators in the data.

For black raters, the estimated population effect size (ρ_{pb}) was $-.220$ and was based on a sample of 14 studies and 2,428 ratees. The 95% confidence interval about this estimate

also excluded zero ($-.41 < \rho < -.03$), indicating that black raters assigned significantly higher ratings to black ratees than to white ratees. The d value for black raters ($-.45$) suggests that the average black ratee received a higher rating than 67% of the white ratees. Although the population variance of effect sizes for black raters ($\sigma_p^2 = .032$) was also greater than zero, moderator analyses were not performed because of the small number of studies available. A Z test by rater race (Rosenthal & Rosen, 1982) showed a highly significant difference in average effect sizes for black and white raters ($z = 7.57$; $p < .001$).

Moderator Analyses for White Ratets

The effect of saliency of blacks on rating differences was tested by computing the correlation across all studies of the uncorrected study effect size and the percentage of black ratees. The correlation was found to be $-.16$ ($N = 73$; $p < .10$) or $-.18$ when corrected for measurement error. The negative sign indicates that effect size increases as the percentage of blacks decreases.

The results of the tests for the four dichotomous moderators are presented in Table 1. Effect sizes found in field setting ($\bar{r}_{pb} = .192$) were significantly larger than effect sizes found in laboratory settings ($\bar{r}_{pb} = .037$; $z = 2.41$; $p < .05$). Because the 95% confidence intervals for laboratory studies includes zero, the hypothesis of no race differences cannot be rejected.

Due to the large discrepancy between laboratory and field results, further tests for moderators were completed with both the entire sample and field studies alone. The results for both sets of studies revealed similar patterns. Theoretically, the results for field studies alone should be more relevant because the remaining moderators (rater training, rater format, and rating purpose) represent organizational conditions rather than psychological processes that may occur in any setting. Therefore, tests of potential moderators for field studies are presented. Examination of the field data revealed no significant differences in subgroup effect sizes for rater training, rating format, or rating purposes. Though nonsignificant, race effects were slightly higher for trained raters than for untrained raters

and for ratings made with behavioral scales versus trait scales.

Discussion

An examination of the results indicates that effect sizes for white raters were positive in 68 of the 74 studies but significantly greater than zero in only 46 cases. A traditional literature review might label such results as conflicting or mixed. The meta-analysis results provide a more definitive conclusion regarding ratee race effects and potential moderator variables.

The size of the ratee race effect is somewhat small for both black and white raters but is relatively consistent across studies. Of greater importance is the difference in the direction of the effect for the two sets of raters. Both black and white raters gave significantly higher ratings to members of their own race. This evidence concurs with statements made by Landy and Farr (1980) and lead us to conclude that race of ratee does have an impact on performance ratings in real world settings. This tendency is equally strong for both black and white raters.

The investigation of potential moderators for white raters demonstrated that effect size was not influenced by type of rating, rating purpose, or rater training. The absence of a moderating effect for type and purpose of rating is contrary to prevalent theory in the performance appraisal domain (e.g., Wherry & Bartlett, 1982). Behaviorally based ratings appear equally prone to race effects as trait ratings. The results are consistent with Landy and Farr's (1980) conclusion that rating formats account for little variance in ratings. Interestingly, Feldman (1981) has noted that the cognitive processes of raters involving the observation, categorization, and explanation of behavior may be impervious to changes in rating format. In other words, raters may selectively attend to and recall behaviors that validate their underlying global (trait) impressions.

The lack of moderating effect for rating purpose was somewhat unexpected because purpose has been found to affect systematic rating effects such as halo and leniency. It appears that the impact of rating purpose does not generalize to ratee race effects. This

reinforces Warmke and Billings's (1979) call for more emphasis on understanding how organizational contextual variables such as rating purpose affect rating criteria.

The lack of an effect for rater training may reflect the fact that rater training programs rarely address issues of differential treatment of ratee subgroups. From this perspective, it could be argued that the data do not allow for an adequate test of the potential impact of rater training on race effects. Alternatively, the results can be viewed as consistent with recent reviews (cf. Bernardin & Pence, 1980) suggesting that rater training has minimal impacts on rating effects. At the very least, the results cast doubt on the rise of traditional rater training programs to reduce ratee race effects in ratings.

A substantial moderating effect was found for the study setting. Race effects, while substantial in field settings, were negligible in laboratory studies. This finding directly contradicts assertions by Wendelken and Inn (1981) and others that race effects are minimized in the field as opposed to laboratory settings. One explanation for this finding is that researchers can ensure equal performance among subgroups in laboratory but not field settings. An alternative (but not contradictory) explanation is that any tendency of subjects to evaluate blacks and whites differentially in laboratory settings is overwhelmed by their desire to give socially desirable responses.

As an illustration, a recent study of sex effects in the rating of essays (Buhrke & Yanico, 1982) found no overall differences in evaluations of male and female writers. Interestingly, when the researchers examined the response of subjects scoring high on a social desirability scale, both male and female subjects assigned higher ratings to female writers.

Regardless of the explanation, the results have implications for the appropriate roles of laboratory and field research. As others have noted (e.g., Berkowitz & Donnerstein, 1982; Mook, 1983), the true value of laboratory research is the explicit examination of well-articulated theoretical propositions and not the estimation of the magnitude of relationships in a population. A good example of the appropriate use of laboratory settings is Schmitt and Lippin's (1980) test of the hy-

pothesis that raters exhibit less confidence and hence have less variability in their ratings of different race or sex ratees. In other words, the theoretical processes or conditions underlying race effects can legitimately be studied in laboratory settings, for these processes may be generalizable to the real world; effect sizes cannot.

The meta-analysis also shows that saliency of blacks in the workgroup may be a moderator. Race effects decline as the percentage of blacks increase. For example, although blacks made up only 22% of the total sample, the percentage of blacks in the 10 studies with the smallest (near zero) effect sizes was 39%. This finding is consistent with theory and research in the social cognition literature. Taylor (1981), for example, has theorized that as the percentage of group members possessing a certain characteristic (race, sex, age, etc.) decreases, that characteristic becomes more salient for processing and recalling information about those particular members. As a result, ratings of those members grow more extreme. The weakness of the present meta-analysis is that the percentage of blacks rated by individual raters was unknown in studies employing multiple workgroups or raters. Moreover, a low percentage does not necessarily imply high saliency. Nonetheless, the results do corroborate predictions from other areas of psychology with real-world data and suggest a fruitful line of research for I/O psychologists.

As noted in the introduction, the results of the present meta-analysis do not directly isolate the effects of racial bias or performance differences in performance evaluations. A closer inspection of the results, though, provides evidence that some portion of rating variance is probably attributable to rater bias. First, raters evaluated same-race ratees higher than different-race ratees. Because the two sets of raters evaluated many of the same ratees, a logical conclusion is that the ratings were biased to some degree. Second, the percentage of blacks in the sample was found to be inversely related to the size of the race effects. This finding suggests that the saliency of a ratee characteristic may be directly related to the degree to which that characteristic is incorporated in performance ratings.

On the other hand, the existence of bias

effects does not preclude the possibility that actual performance differences between races existed. In fact, the data imply that true differences were present in some cases. First, the rater effect size was quite robust under conditions that focused the attention of the rater on actual job behaviors (i.e., behaviorally based rating scales, rather training). From this perspective, the fact that the conditions did not minimize (and tended to increase) the race effect suggests actual performance differences. Second, the effect size was much larger in field than laboratory settings. Although laboratory investigations typically equate subgroup performance, equal subgroup performance distributions in field settings cannot be assumed. Research on selection tests have shown that white applicants tend to score higher on a variety of preemployment ability measures (e.g., see Schmidt, & Hunter, 1981). Therefore, the higher effect sizes for field studies (with white raters) may, in part, be a reflection of the differences found on the predictor measures.

Consequently, no firm conclusions can be reached regarding the extent to which the results found are due to rater bias or ratee performance. Rather, the evidence suggests that differential ratings are more likely due to some combination of bias and performance differences.

More important than isolating the source of the observed differences between races in this meta-analysis is a consideration of the impact of these differences on future research. Landy and Farr (1980), addressing the source of bias, have conceptualized bias as the application of different mental processes of the rater as a function of ratee subgroup. If one views the racial differences found as due to bias, research efforts should focus on the extent to which (a) race is used as a relevant category for observing, storing, and recalling ratee information; and (b) raters differentially weight job- and non-job-relevant factors in evaluating performance of black and white ratees. The social cognition literature provides directions for examining differences in the way in which information is processed by raters. For example, research on tokenism (Kanter, 1977) might help explain why blacks tend to receive more negative ratings when they constitute only a small percentage of the

workforce. In terms of differential weighting of factors, Bass and Turner (1973), Campbell et al. (1973) and others have provided evidence that ratings of black ratees are more related to objective performance measures than white ratees. An interesting follow-up to the present study would be to examine relations between objective and subjective criteria by ratee race using meta-analytic techniques.

The viewpoint that the race effects found in the present study may reflect actual performance differences leads to a research strategy that examines factors in society and organizational settings that foster performance differences. For example, cultural anthropologists such as Sanday (1976) suggest that differences in performance might occur over time because the nondominant culture (e.g., black culture) may not share the common values, perceptions, and institutions of the dominant mainstream culture that defines effectiveness. From an organizational perspective, Ilgen and Youtz (1984) have suggested that lost opportunity factors such as the lack of mentors for blacks, ingroup/outgroup effects (Dansereau, Graen, & Haga, 1975), and self-limiting behaviors perpetuate performance differences between black and white group members.

Consequently, the results of the present study are provocative in directing future research. One research direction is to examine the information processing strategy of the rater, whereas a second direction is to examine contextual factors that might have an impact on individual effectiveness. Based on the present study, both directions seem reasonable and need to be pursued.

In conclusion, this article makes two primary contributions to understanding the impact of race on performance ratings. First, a population effect size has been determined, and this value is the best estimate of what a personnel researcher may expect when collecting ratings. There is little need for future studies that focus exclusively on determining ratee race effect size. Second, the results (particularly of the moderators) should serve to shift the emphasis from examining race effects to a greater understanding of the process underlying race effects. Examination of mean difference in ratings by ratee racial

group does little to increase our understanding of the complex nature of performance differences and racial bias. Instead, we need to develop methods to detect the presence of bias and to investigate the boundaries and covariates of bias in performance evaluations.

References

- Arnold, H. J. (1982). Moderator variables: A clarification of conceptual, analytic, and psychometric issues. *Organizational Behavior and Human Performance*, 29, 143-174.
- Bass, A. R., & Turner, J. N. (1973). Ethnic group differences in relationships among criteria of job performance. *Journal of Applied Psychology*, 57, 101-109.
- Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist*, 37, 245-257.
- Bernardin, H. J. (1978). Effects of rater training on leniency and halo errors in student ratings of instructors. *Journal of Applied Psychology*, 63, 301-308.
- Bernardin, H. J., & Pence, E. (1980). Effects of rater training: Creating new response sets and decreasing accuracy. *Journal of Applied Psychology*, 65, 60-66.
- Bigoness, W. J. (1976). Effect of applicant's sex, race, and performance on employers' performance ratings: Some additional findings. *Journal of Applied Psychology*, 61, 80-84.
- Borman, W. (1979). Format and training effects on rating accuracy and rater errors. *Journal of Applied Psychology*, 64, 410-421.
- Brugnoti, G. A., Campion, J. E., & Basen, J. A. (1979). Racial bias in the use of work samples for personnel selection. *Journal of Applied Psychology*, 64, 119-123.
- Buhrke, R. A., & Yanico, B. J. (1982, August). *Effects of sex-role ideology on sex bias in performance appraisal*. Paper presented at the meeting of the American Psychological Association, Washington, DC.
- Campbell, J. T., Crooks, L. A., Mahoney, M. H., & Rock, D. A. (1973). *An investigation of sources of bias in the prediction of job performance: A six year study (Report PR-73-27)*. Princeton, NJ: Educational Testing Service.
- Cascio, W. F., & Valenzi, E. R. (1978). Relations among criteria of police performance. *Journal of Applied Psychology*, 63, 22-28.
- Dansereau, F., Graen, G., & Haga, W. J. (1975). A vertical dyadic linkage approach to leadership within formal organizations. *Organizational Behavior and Human Performance*, 13, 46-78.
- deJong, J. E., & Kaplan, H. (1962). Some differential effects of race of rater and ratee on early peer ratings of combat aptitude. *Journal of Applied Psychology*, 46, 370-374.
- Farr, J. L., O'Leary, B. S., & Bartlett, C. J. (1971). Ethnic group membership as a moderator of the prediction of job performance. *Personnel Psychology*, 24, 609-636.
- Feldman, J. M. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology*, 66, 127-148.
- Fox, H., & Lefkowitz, J. (1974). Differential validity: Ethnic group as a moderator in predicting job performance. *Personnel Psychology*, 27, 209-223.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 10, 3-8.
- Guion, R. M. (1965). *Personnel testing*. New York: McGraw-Hill.
- Hamner, W. C., Kim, J. S., Baird, L., & Bigoness, W. J. (1974). Race and sex as determinants of ratings by potential employers in a simulated work-sampling task. *Journal of Applied Psychology*, 59, 705-711.
- Hays, W. L. (1973). *Statistics for the social sciences*. New York: Holt, Rinehart, & Winston.
- Hunter, J. E., Schmidt, F. L., & Jackson, G. (1982). *Meta-analysis: Cumulating research findings across studies*. Beverly Hills, CA: Sage.
- Ilgel, D. R., & Youtz, M. (1984, February). *Factors affecting the evaluation and development of minorities in organizations*. Presented at the Office of Naval Research Conference on minorities entering high-technology careers, Pensacola, FL.
- Ivanovich, J. (1979). Longitudinal study of the effects of rater training on psychometric errors in ratings. *Journal of Applied Psychology*, 64, 502-508.
- Jacobs, R., Kafry, D., & Zedeck, S. (1980). Expectations of behaviorally anchored rating scales. *Personnel Psychology*, 33, 595-640.
- Jones, E. E., & Sigall, H. (1971). The Bogus Pipeline: A new paradigm for measuring affect and attitudes. *Psychological Bulletin*, 76, 349-364.
- Kanter, R. M. (1977). *Men and women of the corporation*. New York: Basic Books.
- King, L. M., Hunter, J. E., & Schmidt, F. L. (1980). Halo in multidimensional forced choice performance evaluation scale. *Journal of Applied Psychology*, 65, 507-516.
- Landy, F. J., & Farr, J. L. (1973). *Police performance appraisal*. University Park, PA: Department of Psychology, The Pennsylvania State University.
- Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin*, 87, 72-107.
- Mobley, W. H. (1982). Supervisor and employee race and sex effects on performance appraisals: A field test of adverse impact and generalizability. *Academy of Management Journal*, 25, 598-606.
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 38, 379-387.
- Rosenthal, R., & Rubin, D. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin*, 92, 500-504.
- Sanday, P. R. (1976). Cultural and structural pluralism in the United States. In P. R. Sanday (Ed.), *Anthropology and the public interest: Fieldwork and theory* (pp. 53-73). New York: Academic Press.
- Schmidt, F. L., & Hunter, J. E. (1981). Employment testing: Old theories and new research findings. *American Psychologist*, 36, 1128-1137.
- Schmidt, F. L., Hunter, J. E., & Pearlman, K. (1981). Task differences as moderators of aptitude test validity in selection: A red herring. *Journal of Applied Psychology*, 66, 166-185.
- Schmidt, F. L., Hunter, J. E., & Pearlman, K. (1982). Progress in validity generalization: Comments on Callender and Osburn and further developments. *Journal of Applied Psychology*, 67, 835-845.

- Schmidt, F. L., & Johnson, R. H. (1973). Effect of race on peer ratings in an industrial setting. *Journal of Applied Psychology*, 57, 237-241.
- Schmitt, N., & Hill, T. E. (1977). Sex and race composition of assessment center groups as a determinant of peer and assessor ratings. *Journal of Applied Psychology*, 62, 261-264.
- Schmitt, N., & Lippin, M. (1980). Race and sex as determinants of the mean and variance of performance ratings. *Journal of Applied Psychology*, 65, 428-435.
- Sharon, A. T., & Bartlett, C. J. (1969). Effect of instructional conditions in producing leniency on two types of rating scales. *Personnel Psychology*, 23, 251-263.
- Taylor, E. L., & Wherry, R. J. (1951). A study of leniency in two rating systems. *Personnel Psychology*, 4, 39-47.
- Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 83-114). Hillsdale, NJ: Erlbaum.
- Terborg, J. R., Lee, T. W., Smith, F. J., Davis, G. A., & Turbin, M. S. (1982). Extension of the Schmidt and Hunter validity generalization procedure to the prediction of absenteeism behavior from knowledge of job satisfaction and organizational commitment. *Journal of Applied Psychology*, 67, 440-449.
- Toole, D. L., Gavin, J. F., Murdy, L. B., & Sells, S. B. (1972). The differential validity of personality, personal history, and aptitude data for minority and nonminority employees. *Personnel Psychology*, 25, 661-672.
- Warmke, D. L., & Billings, R. S. (1979). Comparison of training methods for improving the psychometric quality of experimental and administrative performance ratings. *Journal of Applied Psychology*, 64, 124-131.
- Wendelken, D. J., & Inn, A. (1981). Nonperformance influences on performance evaluations: A laboratory phenomenon? *Journal of Applied Psychology*, 66, 149-158.
- Wherry, R. J., Sr., & Bartlett, C. J. (1982). The control of bias in ratings. *Personnel Psychology*, 35, 521-551.

Appendix

Studies Included in the Meta-Analysis of Race Effects

- Arnold, B. C. (1968). Comparison of Caucasian and Negro subgroups on criterion indices of overall job effectiveness. *Dissertation Abstracts International*, 30(2), 881B. (University Microfilms No. 69-12, 499).
- Arvey, R. D., & Gordon, M. E. (1977). *Final report: Validation of test instruments for the identification of successful and unsuccessful police patrol officers*. Knoxville, TN: The University of Tennessee.
- Baehr, M., Saunders, D. R., Froemel, E. C., & Furcon, J. E. (1971). The prediction of performance for black and white police patrolmen. *Professional Psychology*, 2, 46-58.
- Bartlett, C. J., Goldstein, I. L., Mosier, S., Hannan, R., Buxton, V., Simmons, V., Cooper, C. (1977). *An analysis of the validity of the PPA police examination for entry-level selection in the Prince George's Police Department*. College Park, MD: Training and Education Research Programs.
- Bass, A. R., & Turner, J. N. (1973). Ethnic group differences in relationships among criteria of job performance. *Journal of Applied Psychology*, 57, 101-109.
- Bigoness, W. J. (1976). Effect of applicant's sex, race, and performance on employers' performance ratings: Some additional findings. *Journal of Applied Psychology*, 61, 80-84.
- Blumberg, M., Farr, J., Landy, F., Neidig, R., Saal, F., & Whitaker, L. (1974). *Report on Standard Pressed Steel validation project*. University Park, PA: The Pennsylvania State University (Department of Psychology).
- Brugnot, G. A., Campion, J. E., & Basen, J. A. (1979). Racial bias in the use of work samples for personnel selection. *Journal of Applied Psychology*, 64, 119-123.
- Campbell, J. T., Crooks, L. A., Mahoney, M. H., & Rock, D. A. (1973). *An investigation of sources of bias in the prediction of job performance: A six-year study (Report PR-73-27)*. Princeton, NJ: Educational Testing Service.
- Cascio, W. F., & Valenzi, E. R. (1978). Relationships among criteria of police performance. *Journal of Applied Psychology*, 63, 22-28.
- Cox, J. A., & Krumboltz, J. D. (1958). Racial bias in peer ratings of basic airmen. *Sociometry*, 21, 292-299.
- deJong, J. F., & Kaplan, H. (1962). Some differential effects of race of rater and ratee on early peer ratings of combat aptitude. *Journal of Applied Psychology*, 46, 370-374.
- Farr, J. L., O'Leary, B. S., & Bartlett, C. J. (1971). Ethnic group membership as a moderator of the prediction of job performance. *Personnel Psychology*, 24, 609-636.
- Farr, J. L., O'Leary, B. S., Pfeiffer, C. M., Goldstein, I. L., & Bartlett, C. J. (1971). *Ethnic group membership as a moderator in the prediction of job performance: An examination of some less traditional predictors*. Pittsburgh: American Institute for Research.
- Feldman, J. M., & Hiltzman, R. J. (1977). Sources of bias in performance evaluation: Two experiments. *International Journal of Intercultural Relations*, 1, 35-37.
- Field, H. S., Bayley, G. A., & Bayley, S. M. (1977). Employment test validation for minority and nonminority production workers. *Personnel Psychology*, 30, 37-46.
- Ford, K. A. (1975). Ethnic group differences in employment test/job performance relationships. *Dissertation Abstracts International*, 37(9), 739B. (University Microfilms No. 77-24, 829).
- Fox, H., & Lefkowitz, J. (1974). Differential validity: Ethnic group as a moderator in predicting job performance. *Personnel Psychology*, 27, 209-223.

- French, J. (1973). *Development and concurrent validation of selection devices for typist clerks: An interjurisdictional research study*. Lansing, MI: Michigan Department of Civil Service.
- Gael, S., & Grant, D. L. (1972). Employment test validation for minority and nonminority telephone company service representatives. *Journal of Applied Psychology*, 56, 135-139.
- Hakel, M. D., Appelbaum, L., Lyness, K. S., & Moses, J. L. (1982). *Reliable and impartial ratings of management potential*. Unpublished manuscript.
- Hamner, W. C., Kim, J. S., Baird, L., & Bigoness, W. J. (1974). Race and sex as determinants of ratings by potential employers in a simulated work-sampling task. *Journal of Applied Psychology*, 59, 705-711.
- Huck, J. R., & Bray, D. W. (1976). Management assessment center evaluations and subsequent job performance of white and black females. *Personnel Psychology*, 29, 13-30.
- Ivancevich, J. M., & McMahon, J. T. (1977). Black-white differences in a goal-setting program. *Organizational Behavior and Human Performance*, 20, 287-300.
- Kesselman, G. A., & Lopez, F. E. (1979). The impact of job analysis on employment test validation for minority and nonminority accounting personnel. *Personnel Psychology*, 32, 91-108.
- Kirkpatrick, J. J., Ewen, R. B., Barrett, R. S., & Katzell, R. A. (1968). *Testing and fair employment: Fairness and validity of personnel tests for different ethnic groups*. New York: New York University Press.
- Kliesner, C. (1971). *Report on the validity of our para-professional exam series*. Columbus, OH: State of Ohio Personnel Department.
- Kraiger, K. (1981). *Measuring police officer performance: Criterion development for the Columbus Police Officer Selection Validation Project*. Columbus, OH.
- Kriska, S. D., Hines, C. U., & Katko, D. P. (1983). *Criterion-related validity study of the Columbus, Ohio firefighter job*. Columbus, OH.
- Landy, F. J., & Farr, J. L. (1973). *Police performance appraisal*. University Park, PA: The Pennsylvania State University (Department of Psychology).
- Lopez, F. M. (1966). Current problems in test performance of applicants. I. *Personnel Psychology*, 19, 10-18.
- McCann, F. E., Zupkis, R., Howeth, W. F., & Nichols, G. V. (1976). *The validation of the McCann Associates ESV police officer written test*. Huntington Valley, PA: McCann Associates.
- Miller, N., & Angelo, T. (1976). *Concurrent validation study: Income Maintenance Worker I & II*. Columbus, OH: Ohio Department of Administrative Services.
- Mobley, W. H. (1982). Supervisor and employee race and sex effects of performance appraisals: A field study of adverse impact and generalizability. *Academy of Management Journal*, 25, 598-606.
- Moore, M. H. (1973). An investigation of the influence of ethnic group membership on job attitudes and the relationship between these attitudes and job performance. *Dissertation Abstracts International*, 34(4), 1783B-1784B. (University Microfilms No. 73-22, 922).
- Morstein, B. R., & Hsu, L. (1977). *Delaware cooperative police selection study*. Newark, DE: College of Urban Affairs and Public Policy, University of Delaware.
- Neidt, C. Q. (1968). *Report on differential predictive validity of specified selection techniques within designated subgroups of applicants for Civil Service positions*. Colorado State University: Colorado Civil Rights Commission.
- Richards, S. A., & Jaffee, C. L. (1972). Blacks supervising whites: A study of interracial difficulties in working together in a simulated organization. *Journal of Applied Psychology*, 56, 234-240.
- Rosenfeld, M., & Thornton, R. F. (1976). *The development and validation of a firefighter selection examination for the City of Philadelphia*. Princeton, NJ: Educational Testing Service.
- Rosenfeld, M., & Thornton, R. F. (1976). *The development and validation of a multijurisdictional police officer examination*. Princeton, NJ: Educational Testing Service.
- Rosenfeld, M., & Thornton, R. F. (1979). *The development and validation of a police selection examination for the City of Philadelphia*. Princeton, NJ: Educational Testing Service.
- Rotter, N., & Rotter, G. S. (1969). *Race, work performance, and merit ratings: An experimental evaluation*. Paper presented at the Eastern Psychological Association Convention at Philadelphia, PA.
- Schmidt, F. L., & Johnson, R. H. (1973). Effect of race on peer ratings in an industrial setting. *Journal of Applied Psychology*, 57, 237-241.
- Schmitt, N., Merritt, R., Fitzgerald, M., Noe, R. (1981). *Results of the NASSP Assessment Center* (Research Report No. 81-4). East Lansing, MI: Michigan State University.
- Schmitt, N., & Lippin, M. (1980). Race and sex as determinants of the mean and variance of performance ratings. *Journal of Applied Psychology*, 65, 428-435.
- Selection Consulting Center. (1974). *The validation of entry-level firefighter examinations in the states of California and Nevada*. Sacramento, CA: Selection Consulting Center.
- Tenopir, M. L. (1967). *Race and socioeconomic status as moderators in predicting machine-shop training success*. Presented at the 75th Annual Convention of the American Psychological Association, Washington, DC.
- Toole, D. L., Gavin, J. F., Murdy, L. B., & Sells, S. B. (1972). The differential validity of personality, personal history, and aptitude data for minority and nonminority employees. *Personnel Psychology*, 55, 661-672.
- Vance, R. J., & Kraiger, K. (1982). *Research manuscript in progress*. Columbus, OH: The Ohio State University.

Received September 21, 1983

Revision received May 14, 1984 ■